ISSN 0035-6751

INTERNATIONAL REVIEW OF ECONOMICS AND BUSINESS - Volume XXXIX, - n. 1 - January 1992

RIVISTA INTERNAZIONALE DI SCIENZE ECONOMICHE E COMMERCIALI

Anno XXXIX Gennaio 1992 N. 1

Pubblicazione mensile - Sped. in abb. postale, gruppo III/70 Bologna

SOMMARIO

SOTTO GLI AUSPICI DELLA
UNIVERSITÀ COMMERCIALE LUIGI BOCCONI
E DELLA UNIVERSITÀ DEGLI STUDI DI MILANO

CEDAM - CASA EDITRICE DOTT. A. MILANI - PADOVA

12/3/92

RIVISTA INTERNAZIONALE DI SCIENZE ECONOMICHE E COMMERCIALI
(INTERNATIONAL REVIEW OF ECONOMICS AND BUSINESS)

Pubblicazione mensile (A monthly journal). Direzione e Redazione (Editorial Office): Via Teuliè 1, 20136 Milano (Italy), Tel. 02-89409031, C.c. postale 47300207.

Abbonamento 1992 (Subscription 1992): Italia (Italy), Lire 180.000; estero (abroad), Lire 250.000. Collezione completa rilegata 1954-1990, prezzo speciale (Whole bound set of back issues, 1954-1990, special offer price) Lire 2.000.000.

CONDIZIONI DI ABBONAMENTO AI PERIODICI «CEDAM»

L'abbonamento è annuo e si rinnova tacitamente per l'anno successivo se non viene disdetto entro il mese di dicembre, con lettera raccomandata. La semplice reiezione di fascicoli non può essere considerata come disdetta. Il canone di abbonamento deve essere pagato anticipatamente. In caso contrario la Casa si riserva la facoltà di interrompere l'invio dei fascicoli. I pagamenti devono essere effettuati direttamente alla Casa di Padova sul c/c postale n. 205351 oppure ai suoi incaricati muniti di speciale delega, che rilasceranno ricevuta sui moduli recanti il marchio Cedam e numerati progressivamente. Il rinnovo dell'abbonamento deve essere effettuato entro il 31 maggio di ogni anno. Trascorso tale termine l'amministrazione provvederà direttamente all'incasso mediante emissione di fattura con ricevuta bancaria. I fascicoli non pervenuti all'abbonato devono essere reclamati prima della conclusione dell'abbonamento in corso. Decorso tale termine saranno spediti, se disponibili, contro rimessa dell'importo. L'abbonamento importa, agli effetti legali, elezione di domicilio in Padova presso la Casa Editrice.

Direttore responsabile: Aldo Montesano - Autorizz. Tribunale di Treviso N. 113 del 22-10-54

USPI Rivista associata all'Unione della Stampa Periodica Italiana

Tip. Leonelli - Villanova di Castenaso (Bo)

Rivista Internazionale di Scienze Economiche e Commerciali
Volume 39 (1992), N. 1, 1-4

L'ETICA IN ECONOMIA PASSA PER L'EFFICIENZA

di
Anna Pellanda *

Si dice che se fallisce il mercato, l'errore s'annida nella meritocrazia che distribuisce solo a chi è capace di produrre. Ma questa critica risale all'Ottocento, a J.S. Mill che ne scrive nel 1848. Essa si indirizza al sistema di mercato (puro) chiamato comunemente capitalismo da chi crede che i fattori di produzione caratterizzino a turno i sistemi economici, come se esistesse un'economia della terra, una del lavoro, una dell'imprenditorialità, ecc. Chi usa cosí superficialmente la terminologia economica dimentica che i fattori di produzione sono tra loro complementari e che tra tutti è sempre presente il lavoro, con ruolo di fattore "originario" e imprescindibile. Ma si sa che quando una scienza è sociale il suo linguaggio non gode della rispettosa esclusiva degli addetti ai lavori e deve rassegnarsi a subire le storpiature di tutti i suoi fruitori.

Queste improprietà non si limitano al vocabolario economico ma si rivestono di vetustà partigiana, come si accennava in esordio, perché trascurano quello che la critica ottocentesca e la realtà economica hanno sollecitato: la nascita dell'economia del benessere. Questa, prima della piú famosa e dirompente teoria keynesiana, indica che se l'economia di mercato è solo efficiente alla Pareto, per non risultare iniqua distributivamente va affiancata dall'attività fiscale e legislativa dello Stato. Ciò è quanto la politica economica ha recepito e attuato almeno nei paesi europei a cosiddetta economia mista. Parlare di mercato puro è oggigiorno quantomeno ingenuo; criticarne la mancata equità distributiva è come minimo inattuale. Ormai è acquisito che non è il mercato che deve darsi carico del benessere collettivo ma lo Stato con il suo apparato legislativo che disciplina proprio le forme di mercato, dalla concorrenza (anch'essa « impura ») con le leggi anti-trust, al monopolio con la fissazione di tariffe nel caso esso sia pubblico. Ed è lo Stato che con il prelievo fiscale deve preoccuparsi di fornire quello che il mercato, per sua logica intrinseca, non può offrire, cioè i beni pubblici. Si precisa subito che è il prelievo fiscale la cornice operativa della spesa pubblica non il "deficit spending" di keynesiana memoria, perché

---

* Università di Padova, Facoltà di Giurisprudenza, Padova.

questo innesca la spirale del debito pubblico rinnovato-per-ripagarsi che danna, almeno in Italia, alla peggiore iniquità tutto il sistema economico.

Stranamente però gli etici censori dell'attività economica indirizzano i loro strali al mercato, non allo Stato. Essi parlano delle ingiustizie del primo tacendo quelle del secondo. A dire il vero non sono quasi mai economisti: infatti la professione d'estrazione neo-classica non ha bisogno di parlare avendo dalla sua la Storia e lo sfacelo rovinoso dei sistemi marxisti; mentre ai neo-keynesiani conviene il silenzio perché hanno esaltato proprio la forma piú subdola d'iniquità sociale e il *"mea culpa"*, si sa, è oneroso per tutti. Entrambi però sanno che in Italia sono il debito dello Stato, l'improduttività della spesa pubblica, il parassitismo della burocrazia da comunale a statale, la distorsione del risparmio privato letteralmente ipnotizzata dalla finanza pubblica, lo spreco di risorse e l'ignoranza delle preferenze (ma bisognerebbe ormai chiamarle "grida di dolore") degli utenti di beni e servizi prodotti dallo Stato che testimoniano la mancanza di Etica in Economia.

Sarebbe tempo di parlare di "fallimento dello Stato" intervenuto originariamente a sanare il "fallimento del mercato" mediante la fornitura di beni pubblici. Dell'utilità di questi (che Smith voleva limitati a giustizia, difesa e pochissime opere pubbliche) pare non si abbia coscienza diretta e questo consente di svicolare dalla teoria economica del valore per flirtare con le teorie oligarchiche, volontaristiche, piú o meno etiche della Finanza pubblica. Si dice inoltre che siano non divisibili (pro-capite), non escludibili (come il loro contrario, i mali pubblici) e che non abbiano prezzo soddisfacendo i bisogni prima di essere recepiti. Sono quindi "fuori" della logica di mercato, tanto è vero che sono venduti a tariffe (non prezzi) spesso inferiori ai costi e affidati a colletti bianchi non solo perché pertengono quasi sempre al terziario ma perché esigono protezione sindacale e monopolistica "da camice" non "da tuta". Questo provoca rivalità distributive specie da parte del secondario e del terziario privato. In realtà tutta questa problematica è distributiva e demandata allo Stato. Ma di Etica della distribuzione, di efficienza dello Stato non si sente tanto parlare tranne che per strada ove s'incolpa della pioggia il governo ma ove, con pari emotività, si profetizzano calamità per inidentificati "padroni".

Ora se l'Etica è rivolta al buono, come l'Estetica lo è al bello, il Diritto all'equo, non bisogna dimenticare che all'Economia pertiene l'utile. Se questo requisito non va travisato, l'utile non deve diventare (se non a discrezione strettamente privata) beneficienza, la distribuzione non deve trasformarsi in assistenzialismo, la professionalità scadere a parassitismo. Ma soprattutto è la produzione che non può sottrarsi all'adozione delle tecniche piú convenienti, al calcolo dei costi rapportati ai ricavi, al soddisfacimento della domanda dei consumatori. Soltanto se questa logica viene rispettata si produce reddito che poi lo Stato può tassare e redistribuire sotto forma di spesa pubblica e quindi creazione di ulteriore reddito, occupazione, ecc. Se non si produce reddito, su che cosa può far leva l'imposizione fiscale? Se poi lo Stato non lo sa spendere produttivamente come può esso intervenire a sanare le insufficienze del merca-

to? La fornitura di beni e servizi pubblici ha una matrice fiscale e una operatività economica. La prima segue criteri d'equità distributiva, la seconda canoni d'efficienza produttiva. Essi si alimentano a vicenda e poggiano entrambi sull'Etica dell'utile.

Ma purtroppo questo non si verifica nel nostro Paese dove a fronte di imprese private i cui imprenditori si confrontano costantemente con la concorrenza interna e internazionale e i lavoratori con lo spettro della disoccupazione, dilaga il terziario pubblico ove i dirigenti devono continuamente affrancarsi dal clientelismo politico e gli impiegati, forti dell'impossibilità legale di licenziamento, si fanno sommergere dal parassitismo. Lo Stato in Italia non sa né ridistribuire né produrre ed è tanto piú iniquo quanto piú inefficiente. Come si può uscire da questa situazione? Mirando al cuore del problema cioè non all'incapacità di uno Stato impersonale quanto a quella dei suoi uomini. Se questi sono politici, un paese democratico li cambia. Se questi sono burocrati, li educa. L'esempio può venire dal Giappone dove l'istruzione svolge un ruolo essenziale nel preparare gli individui alla carriera.

Il sistema educativo giapponese è uno dei "piú monomaniacalmente meritocratici esistenti al mondo" sostiene il professor Ronald Dore; esso soppianta l'elitarismo sociale perché aperto a tutti ed elimina il pressapochismo professionale perché seleziona dall'età della scuola dell'obbligo, cioè dai quindici anni. In Giappone si istruisce per affrontare qualsiasi mestiere incoraggiando l'iniziativa privata per acquisire capacità e abilità lavorative e divulgando i sussidi di Stato perché si vaglino, a livello nazionale e standardizzato, questi raggiungimenti. Con questo sistema si formano operai e impiegati perché i giovani solo nella percentuale del 5 o 6 per cento non proseguono l'istruzione presso qualche scuola superiore oltre il quindicesimo anno. E cosí si preparano e distinguono i migliori che continuano gli studi nelle università, scegliendo queste e venendo da queste scelti, in base alla loro intelligenza naturale e alle loro doti di volontà morale. Una volta finita l'università, i piú preparati ambiscono ad entrare nella pubblica amministrazione essendo questo il settore lavorativo piú qualificato; esso infatti riscuote grande consenso nel paese che sa di avere un apparato burocratico efficiente perché preparato e competente.

Se la concezione nipponica della dedizione individuale e della gratificazione spirituale è di difficile esportazione nel nostro mondo tanto piú materialistico quanto piú la sua Etica mira alla ricompensa oltremondana, uno strumento universalmente valido e applicabile è un sistema educativo severo e diffuso. Solo chi sa fare è efficiente e morale e tutti possono venir istruiti per diventarlo. (agosto 1991)

ETHICS IN ECONOMICS MEANS EFFICIENCY

Ethical criticisms against economic systems are usually addressed to market rules rather than to public sectors. The justification for the State intervention in the economy is the production of public goods; the basis on which the State can provide these goods is, according to us, fiscal policy not the keynesian deficit spending. But in order to have incomes to be taxed, people must produce them. Creating incomes means producing according to the market rules of profit and demand.

The Italian system is disgracefully iniquitous and inefficient because both its fiscal policy is unfair and its productivity is very low. Trying to point out a solution, we propose to consider its men instead of its impersonal character of public entity. Looking at Japan where bureaucrats are competent and efficient because scholarly well prepared, we think that strict schools for public employees could lead to eliminate fiscal and productive inefficiency and thus reach ethical standards.

Rivista Internazionale di Scienze Economiche e Commerciali
Volume 39 (1992), N. 1, 5-22

THE BANG-BANG PRODUCTION OF DEPLETABLE NATURAL RESOURCES

by
DAVID B. REISTER * and MICHAEL A.S. GUTH **

I. *Introduction*

Since the seminal paper by Hotelling (1931), economists have known that when the market is in equilibrium, the net price to an owner of a depletable natural resource must increase with the interest rate. To quote Solow (1974), "It is hard to overemphasize the importance of this tilt in the time profile for net price. If the net price were to rise too slowly, production would be pushed nearer in time and the resource would be exhausted quickly, precisely because no one would wish to hold resources in the ground and earn less than the going rate of return. If the net price were to rise too fast, resource deposits would be an excellent way to hold wealth, and owners would delay production while they enjoyed supernormal capital gains".

However, the behavior of the oil market since 1974 illustrates that the net price of a depletable natural resource can increase faster or slower than the interest rate. Since 1974, a small group of producers have caused two sharp increases in oil prices and one sharp decrease. The oil market has not been in equilibrium and most producers have faced exogenous prices.

In this paper, we calculate the optimal production path for an owner of

* Oak Ridge National Laboratory, Oak Ridge, TN (U.S.A.).
** Credit Suisse First Boston, Ltd., London (Great Britain).

This manuscript was prepared for the Oak Ridge National Laboratory, Oak Ridge, Tennessee, 37831, USA, operated by Martin Marietta Energy Systems, Inc., for the U.S. Department of Energy under Contract No. DE-AC05-840R21400. Financial support for this research was provided by the Office of Planning and Environment of the U.S. Department of Energy. The authors want to thank R.G. Alsmiller, Jr., Dan Christiansen, Bob Conrad, and Vladimir Protopopescu for helpful discussions, comments, and suggestions.

a depletable natural resource for the case where the market is not in equilibrium and the net price is not increasing with the interest rate. We shall begin with the case where the extraction costs are constant and subsequently consider the case where the extraction costs increase with cumulative consumption. We shall find that the Maximum Principle of Pontryagin (1962) is a fruitful method for solving the problems. When the price of the resource is exogenous, the optimal production path is bang-bang; that is, the resource owner is either at full production or at zero production. The key decision for the resource owner is the switch time, when to start or stop production. For our problem, the Hotelling rule is the switching rule, rather than a forecast of the net price. Our results are an extension of the work of Clark (1976).

II. *The Basic Problem*

Consider an owner of a finite stock of a depletable natural resource who knows the future price ($P(t)$) for the resource and wishes to maximize his profits. If his extraction costs are C, then his profit is $P(t) - C$. If the owner uses a discount rate (r) to compare future profits to present profits, then the objective of the resource owner is to maximize the discounted value of his profits (J):

$$J = \int_0^T [P(t) - C]\, q(t)\, e^{-rt}\, dt, \tag{1}$$

where $q(t)$ is the production rate for the resource.

We will assume that the production rate is bounded by a capacity constraint:

$$q_{\max} \geq q(t) \geq 0, \tag{2}$$

where $q_{\max}$ is given and finite. Since the resource is exhaustible, we assume that the owner's stock of the resource is finite:

$$Q(T) = \int_0^T q(t)\, dt \leq Q^*. \tag{3}$$

The optimization problem for the resource owner is to find a production rate [$q(t)$] that satisfies the conditions of Eqs. (2) and (3) and maximizes Eq. (1). We shall call this optimization problem the basic problem.

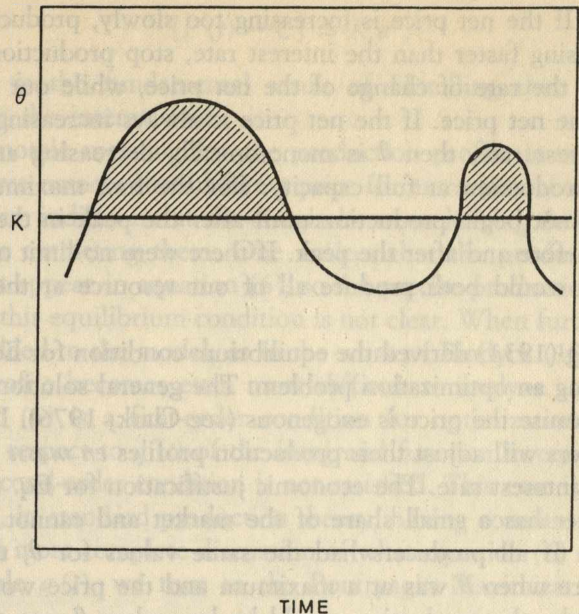

FIGURE 1. The discounted present value of the net price for a depletable natural resource.

We can obtain the solution of the basic problem from first principles. Let θ be the discounted present value of the net price:

$$\theta(t) = [P(t) - C]\, e^{-rt}. \tag{4}$$

Consider the case where θ has the values displayed in Fig. 1; that is, θ has two maxima and the first is larger than the second.

When should an owner of a finite stock of a depletable natural resource sell the resource? He should sell the first unit when θ is at a maximum and he should sell additional units near the maximum until he depletes his stock. If his stock of resource is large enough, he can produce during both maxima. For the case displayed in Fig. 1, the resource owner has the following bang-bang decision rule: produce at full capacity whenever θ is greater than K and stop production whenever θ is less than K.

For the values of $\theta(t)$ plotted in Fig. 1, each value of K between the maximum value of θ and zero is the solution of the basic problem for a mix of demand $[D(t)]$ and total stock of the resource $[Q^*]$. If the resource owner has a small stock, then K should be near the maximum value of θ. If the owner has a large stock, then K can be near zero.

The quote from Solow (1974) suggests the following bang-bang production rule: If the net price is increasing too slowly, produce now. If the net price is rising faster than the interest rate, stop production. The Solow rule concerns the rate of change of the net price, while our rule concerns the level of the net price. If the net price is always increasing more slowly than the interest rate, then θ is monotonically decreasing and both rules recommend production at full capacity. For the first maximum in Fig. 1, Solow would not begin production until after the peak in the curve, while we produce before and after the peak. If there were no limit on the production rate, we would both produce all of our resource at the peak of the curve.

Hotelling (1931) derived the equilibrium condition for Eq. (1) without formally posing an optimization problem. The general solution of Eq. (1) is bang-bang because the price is exogenous (see Clark, 1976). In equilibrium resource owners will adjust their production profiles *en masse* so that prices rise with the interest rate. The economic justification for Eq. (1) is that the resource owner has a small share of the market and cannot influence the market price. If all producers had the same values for θ, then everyone would produce when θ was at a maximum and the price would be driven down. Similarly, the production would be low when θ was at a minimum and the price would be driven up.

Although market equilibrium may require that the net price increase with the interest rate, the time delays inherent in discovering and developing a depletable natural resource may prevent the market from ever reaching equilibrium. For the oil market, the price in 1981 was probably too high, while the price in 1986 was probably too low.

Because the price is exogenous, the optimum solution to the basic problem is bang-bang; when the price is right, the resource owner produces as much as possible. The limit on production (Eq. (2)) is required to guarantee a finite solution. How realistic is the limit on production? Most mines or wells have an upper limit on production capacity. In most cases, mines or wells are designed to operate for several years rather than for days or weeks. In a later section, we will assume that the price depends on the production rate and time.

We can attempt to solve the basic problem using the Calculus of Variations (see Sagan, 1969). Applying the Euler-Lagrange equation to the problem, the partial derivative of the integrand of Eq. (1) with respect to $q\,(t)$ is a constant:

$$[P(t) - C]\,e^{-rt} = K, \tag{5}$$

where K is a constant. Equation (5) may be rewritten:

$$[P(t) - C] = K e^{rt}, \tag{6}$$

Equation (6) is the fundamental result of Hotelling that the net price increases with the interest rate.

The economic meaning for the production profile in equation (6) is that all resource owners should face a price for extraction of their resources that leaves them indifferent at the margin between leaving their resources in the ground or extracting them. One notices that the production variable, $q(t)$, does not appear in equation (6), so the precise production profile that accomplishes this equilibrium condition is not clear. When further stochastic variation is added to the model, as in the work of Pindyck (1981), then the production profile becomes even more difficult to analyze.

Equation (6) is a first-order condition for optimization. Due to the linearity with respect to $q(t)$ of the integrand function in equation (1), the (sufficient) second-order condition is not satisfied. This case is analogous to that of a firm, in standard production theory, having constant average cost. To allow for a more complete discussion of what equation (6) means for the control variable $q(t)$, we turn to the Pontryagin Maximization Principle.

III. *The Pontryagin Maximum Principle*

In this section, we will briefly introduce the maximum principle. A full and rigorous presentation may be found in Pontryagin (1962). Consider the optimal control problem of finding a control vector $[u(t)]$ that will move an object from one point in state space $[x(0)]$ to another point $[x(T)]$ and minimize a functional (J):

$$J = \int_0^T f^0[x(t), u(t)]\, dt. \tag{7}$$

The laws of motion for the object can be written in the form of a system of differential equations:

$$\frac{dx^i}{dt} = f^i[x, u], \quad \text{for} \quad i = 1, \ldots, n. \tag{8}$$

Note that the laws of motion and the integrand of the objective function are autonomous; that is, they do not depend explicitly on time.

To solve the problem, we introduce a system of auxiliary variables $[\psi]$ that satisfy the following equations:

$$\frac{d\psi_i}{dt} = - \sum_{j=0}^{n} \frac{\partial f^j}{\partial x^i} \psi_j, \quad \text{for} \quad i = 0, 1, ..., n. \tag{9}$$

Using the auxiliary variables, we define a Hamiltonian function [H] by:

$$H[\psi, x, u] = \sum_{j=0}^{n} \psi_j f^j [x, u]. \tag{10}$$

We shall say that $u(t)$ is an admissible control if it is piecewise continuous for $0 \leq t \leq T$ and its range is in a set U. Let M be the least upper bound of the Hamiltonian with respect to u:

$$M[\psi, x] = \sup_{u \in U} H[\psi, x, u]. \tag{11}$$

The Pontryagin Maximum Principle. – Let u be an admissible control. Then u is an optimal control if:

1. u maximizes H; that is, $H[\psi, x, u] = M[\psi, x]$, and
2. at the terminal time (T), $\psi_0(T) \leq 0$ and $M[\psi(T), x(T)] = 0$.

Furthermore, ψ_0 and $M[\psi, x]$ are constant.

To solve an optimal control problem using the Pontryagin Maximum Principle, we define the auxiliary variables and find the control that maximizes the Hamiltonian.

IV. *Solution of the Basic Problem Using the Maximum Principle*

For the basic problem, the control variable is the production rate, $q(t)$. The first state variable [x_1] is the cumulative production, $Q(t)$. The Maximum Principle requires that the integrand of the objective function be autonomous. To make the integrand autonomous, we introduce time as a second state variable, x_2.

To summarize the basic problem, the components of the function f are given by:

$$f^0 = -\{P[x_2] - C\} u \exp[-rx_2], \tag{12}$$

$$f^1 = u, \text{ and} \tag{13}$$

$$f^2 = 1. \tag{14}$$

Since ψ_0 is a negative constant and the system of auxiliary equations is linear and homogeneous, we can make an arbitrary choice for ψ_0; let $\psi_0 = -1$. For the basic problem, the Hamiltonian function may be written:

$$H[\psi, x, u] = -f^0 + u\,\psi_1 + \psi_2. \tag{15}$$

The optimal control maximizes the Hamiltonian. Since f^2 does not depend on u, the second auxiliary variable [ψ_2] does not influence the solution and we will ignore it.

The first auxiliary variable satisfies the following equation:

$$\frac{d\psi_1}{dt} = \frac{\partial f^0}{\partial x_1}. \tag{16}$$

For the basic problem, the extraction costs do not depend on cumulative production. Thus, the right side of Eq. (16) is zero and the first auxiliary variable is a constant. Later, we shall allow the extraction costs to depend on cumulative production.

If we rewrite the Hamiltonian in the original variables and let the first auxiliary variable equal $-K$, the Hamiltonian function may be written:

$$H = \phi\, q + \psi_2 \tag{17}$$

where ϕ is defined by:

$$\phi(t) = [P(t) - C]\, e^{-rt} - K. \tag{18}$$

The optimal production rate [q] maximizes the Hamiltonian function. When ϕ is positive then q is at its upper bound, $q = D(t)$; and when ϕ is negative then q is at its lower bound, $q = 0$. In the jargon of optimal control theory, the optimal control is a bang-bang solution.

When we applied the Euler-Lagrange equation to the basic problem, we derived the condition that $\phi = 0$. Now in applying the Pontryagin Maximum Principle to the basic problem, we find that $\phi = 0$ is not the solution: it is the condition for starting or stopping production. The Maximum Principle has provided us with a switching rule to determine the production rate. A more comprehensive discussion of the bang-bang solution to the basic problem and of the application of the Pontryagin Maximum Principle to the optimal management of renewable and nonrenewable resources may be found in Clark (1976).

V. *The General Problem*

In this section, we will apply the Maximum Principle to a more general problem, for which the extraction cost increases with cumulative production and demand shifts over time. We will assume that the extraction cost depends on both cumulative production (Q) and time:

$$C = C[Q, t]. \tag{19}$$

The extraction cost depends on time because changes in technology can reduce production costs.

A basic economic principle is that sales depend on the interplay between supply and demand. The resource owner offers to sell his resource at a price and the market determines the quantity of resource that he will sell:

$$q = F[P, t]. \tag{20}$$

Both Hotelling (1931) and Stiglitz (1976) have considered the optimum production strategy for a monopolist. A monopolist controls the price by setting the level of production:

$$P = P[q, t]. \tag{21}$$

If the functions are single valued, an inverse function exists and there is no mathematical difference between Eqs. (20) and (21). Since we started with q as the control variable, we will continue with q as the control variable. However, we will consider both cases: competition and monopoly.

We have modified the integrand of the basic problem and must redefine the function f^0. The function f^1, f^2, and the Hamiltonian function are unchanged. For the general problem the function f^0 is given by:

$$f^0 = -\{P[u, x_2] - C[x_1, x_2]\} u \exp[-rx_2]. \tag{22}$$

Using the original variables and Eq. (16), the first auxiliary variable satisfies the following equation:

$$\frac{d\psi_1}{dt} = \frac{\partial C}{\partial Q} q e^{-rt}. \tag{23}$$

If we assume that the partial derivative of the extraction cost with respect to cumulative production is positive, the first auxiliary variable increases whenever the production rate is positive. If we assume that the initial value of ψ_1 is $\psi_1 = -K$, then the magnitude of ψ_1 decreases whenever q is positive. At the terminal time, either $\psi_1(T)$ is zero and $Q < Q^*$ or $\psi_1(T)$ is negative and $Q = Q^*$.

The optimal production rate maximizes the Hamiltonian. To find the optimal production rate, we differentiate the Hamiltonian with respect to q; the result is:

$$\{P[q, t] - C[Q, t] + \frac{\partial P}{\partial q} q\} e^{-rt} = -\psi_1 . \tag{24}$$

Equation (24) is the solution to the general problem. Given a demand function $\{P[q, t]\}$ and a production cost function $\{C[Q, t]\}$, Eq. (23) can be solved to determine ψ_1 and Eq. (24) can be solved to determine the production rate $[q]$.

Hotelling (1931) used the Euler-Lagrange equation to derive Eq. (24) for the case of constant extraction costs. When the extraction costs are constant, Eq. (24) states that the marginal revenue minus the production cost increases at the interest rate (see Gordon, 1967). We are not aware of any previous derivation of Eq. (23). For the general problem, the price depends on the production rate and the Euler-Lagrange equation can be used to solve the problem. The Pontryagin Maximum Principle illuminates the solution by introducing the auxiliary variables.

Stiglitz (1976) has derived Eq. (24) for the special case of a constant elasticity of demand and extraction costs that depend on time but not on cumulative production:

$$P[q, t] = h(t)\, q^{\alpha-1}, \tag{25}$$

and

$$C[Q, t] = g(t) . \tag{26}$$

Weinstein and Zeckhauser (1975) have derived a result similar to Eq. (23) for a discrete time problem with increasing production costs. However, they do not define the auxiliary variable.

If the Euler-Lagrange equation is applied to the general problem, the condition for optimality is the equation that results when Eq. (24) is differentiated with respect to time and Eq. (23) is used to eliminate the auxiliary variable. The optimality condition without the auxiliary variable has been derived for the renewable resource problem by Clark (1976), Clark and Munro (1975), Berck (1981), and Pindyck (1984).

We can make Eq. (24) more understandable by defining an exhaustible resource owner's scarcity rent $[R]$ by:

$$R = -\psi_1\, e^{+rt}. \tag{27}$$

We define an objective function $[L]$ by:

$$L = \{P[q, t] - C[Q, t] - R\}\, q. \tag{28}$$

If the resource owner chooses a production rate that maximizes L for each time period, then he will satisfy Eq. (24) and solve the general problem. In the objective function [L], a rent has been added to the extraction costs. The rent term summarizes the dynamics of the problem and incorporates the increases in extraction costs. The rent converts a multi-period optimization problem into a series of single period optimization problems.

VI. *A Logit Demand Function*

To proceed further, we must define a demand function. We assume that the owner sells the resource in a market where a competing resource is offered at price $W(t)$. If the owner's price [$P(t)$] is greater than W, he will lose market share and vice versa. We consider a simple logit demand model:

$$q = D(t)\, s(t), \tag{29}$$

where $D(t)$ is the total demand, and the market share for the resource owner [s] is given by:

$$s(t) = \frac{P^{\gamma}}{P^{\gamma} + W^{\gamma}}, \tag{30}$$

where γ is a parameter. The logit share function (Eq. (30)) has been widely used in models of energy supply and demand; see Boyd, Phillips, and Regulinski (1982) and Reister (1983). Our logit demand model could be used to simulate whether a country uses domestic or imported oil.

Let σ be the price elasticity of demand:

$$\sigma = \frac{P\,\partial q}{q\,\partial P}. \tag{31}$$

Using the definitions of σ and R, Eq. (24) may be written:

$$P\{1 + 1/\sigma\} = C + R. \tag{32}$$

For the logit demand function,

$$\sigma = \gamma\{1 - s\}. \tag{33}$$

Define p and c by: $p = P/W$ and $c = [C + R]/W$. Using the dimensionless parameters p and c, Eq. (32) may be written:

$$c = p\{1 + 1/\sigma\} = G(p). \tag{34}$$

Given c, we would like to find p. Since σ is function of p, $G(p)$ is the inverse function. If we construct a table of $G(p)$ as a function of p, then we can use the table to determine p as a function of c.

The logit exponent $[\gamma]$ controls the price elasticity of the demand model (see Eq. (33)). In many applications in economics, a price elasticity of -2 is a large value. However, if $\gamma = -2$ and the owner's price was 10% higher than the competing price, the resource owner would capture 45% of the market. If the customers are choosing the least cost option, the market share for the more expensive resource would be zero. To reduce the market share for the expensive resource, we will raise the logit exponent to $\gamma = -40$. The functions $G(p)$, $s(p)$, and $\sigma(p)$ are displayed in Table 1 for $\gamma = -40$.

If p is less than 0.91, then $G(p)$ is negative. If p is greater than 0.92, then $G(p)$ is positive. As p increases from 0.92, $G(p)$ increases, $s(p)$

TABLE 1.

THE PRICE-COST FUNCTION AND THE MARKET SHARES FOR THE LOGIT DEMAND FUNCTION

Gamma = − 40.0

p	$G(p)$	$s(p)$	$\sigma(p)$
0.90	− 0.64	0.99	− 0.58
0.91	− 0.10	0.98	− 0.90
0.92	0.25	0.97	− 1.38
0.93	0.48	0.95	− 2.08
0.94	0.64	0.92	− 3.11
0.95	0.74	0.89	− 4.56
0.96	0.81	0.84	− 6.54
0.97	0.86	0.77	− 9.13
0.98	0.90	0.69	− 12.33
0.99	0.93	0.60	− 16.03
1.00	0.95	0.50	− 20.00
1.01	0.97	0.40	− 23.93
1.02	0.98	0.31	− 27.53
1.03	1.00	0.23	− 30.61
1.04	1.01	0.17	− 33.10
1.05	1.02	0.12	− 35.02
1.06	1.03	0.09	− 36.46
1.07	1.04	0.06	− 37.50
1.08	1.05	0.04	− 38.24
1.09	1.06	0.03	− 38.77
1.10	1.07	0.02	− 39.14

decreases, and $\sigma(p)$ becomes more negative. For large values of p, $G(p)$ approaches p, $s(p)$ approaches zero, and $\sigma(p)$ approaches -40.

Given the extraction cost, rent, and competing price, we can calculate c and determine p and s from Table 1. The price ratio [p] is plotted in Fig. 2,

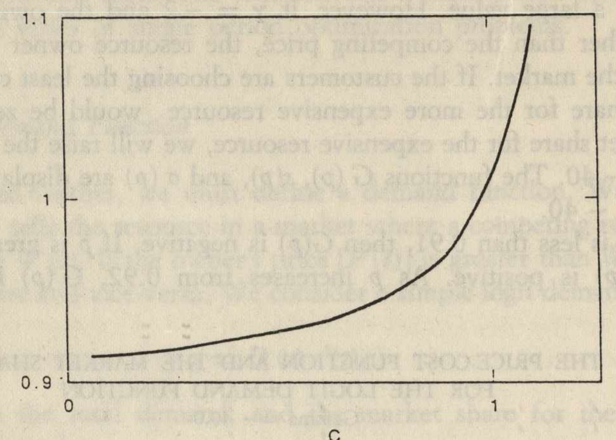

FIGURE 2. The price ratio (p) as a function of the parameter c.

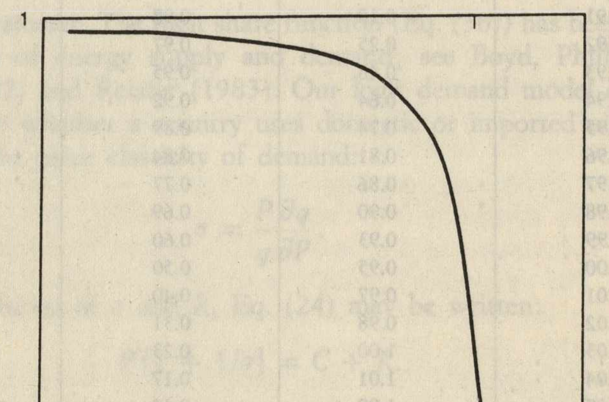

FIGURE 3. The market share (s) as a function of the parameter c.

while the market share [s] is plotted in Fig. 3. For all values of c, p is in the neighborhood of 1.0. Thus, the owner's price for the resource is always close to the competing price. If the competing price declines, the price offered by the owner will decline until $c = 1$. If c is small, p is below 1.0 and the market share is near 100%. Thus, if the sum of the extraction cost and the rent is less than the competing price, the optimum strategy is to have a large market share. If c is greater than 1.0, then p is greater than c and the market share is small. If the competing price falls below the sum of the extraction cost and the rent, the optimum strategy is to have a small market share.

The relationship between the total resource [Q^*] and the initial value for ψ_1 [$\psi_1(0) = -K$] must be determined by numerical integration of Eq. (23). Given the total demand [D], the competing price [W], and the extraction cost function [Q], there will be a range of values for K that determine optimum solutions for various values of Q^*. In general, both the basic problem and the general problem with the logit demand function have bang-bang solutions and the auxiliary variable [ψ_1] controls the starting and stopping of production.

The solution of the general problem is illustrated in Figs 4 through 7. For the example, the price of the competing resource [W] starts at \$ 40, declines to \$ 20 in year 10, before increasing to \$ 60 in year 30. Because the discount rate is 1%, the maximum value for the discounted value of W occurs at the end of the period; that is, the discounted value of W is \$ 45 in year 30. For the example, the total demand is $D = 4$ per year and the maximum production in 30 years is $Q = 120$. The extraction cost has a

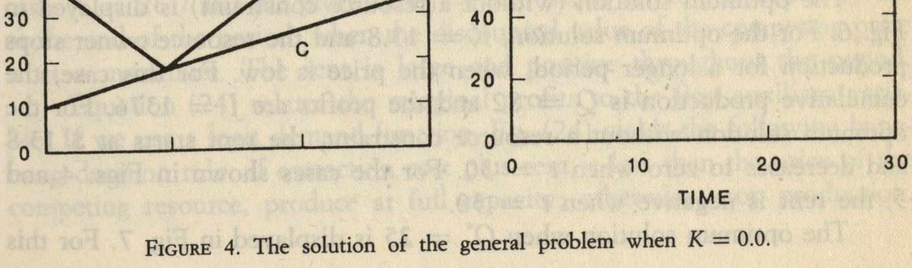

FIGURE 4. The solution of the general problem when $K = 0.0$.

linear dependence on Q : $C = 10 + 0.2\ Q$. As Q increases from 0 to 120, the cost increases from 10 to 34. For each value of K (the initial value for the rent), the equations can be solved to determine the rent, production rate, and discounted profits [J]. By varying K, the maximum value for the profits can be found.

The solution for $K = 0$ is displayed in Fig. 4. When $K = 0$, the rent is negative and the resource owner is near full production for the entire period. For the case displayed in Fig. 4, the cumulative production is $Q = 119$ and the profits are $J = 1094$.

The solution when $K = 10$ is displayed in Fig. 5. For this case, the

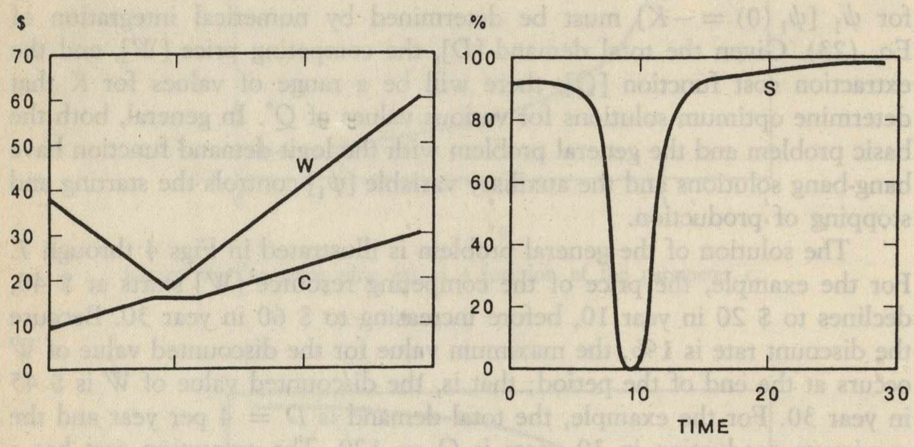

FIGURE 5. The solution of the general problem when $K = 10.0$.

resource owner starts near full production; stops production for a few years when the competing price is low; and returns to full production when the price recovers. By stopping production, the resource owner reduces his costs and increases his profits in the later period. For this case, the cumulative production is $Q = 100$ and the profits are $J = 1312$.

The optimum solution (without a resource constraint) is displayed in Fig. 6. For the optimum solution, $K = 13.8$ and the resource owner stops production for a longer period, when the price is low. For this case, the cumulative production is $Q = 82$ and the profits are $J = 1376$. For the optimum solution without a resource constraint, the rent starts at \$ 13.8 and decreases to zero when $t = 30$. For the cases shown in Figs. 4 and 5, the rent is negative when $t = 30$.

The optimum solution when $Q^* = 25$ is displayed in Fig. 7. For this

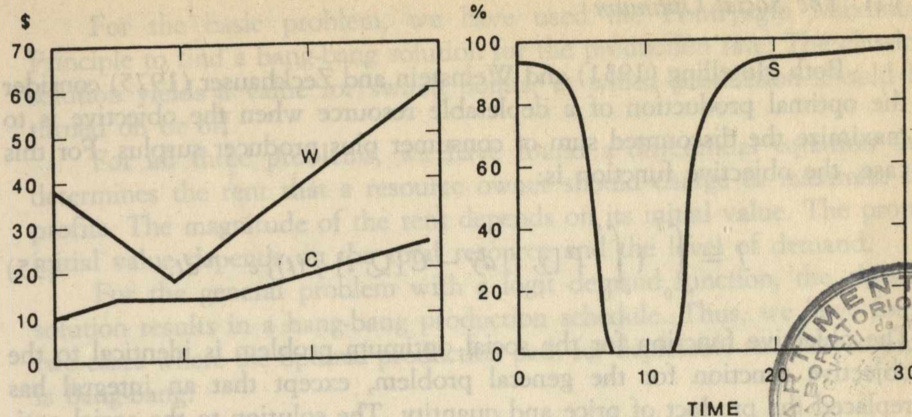

FIGURE 6. The optimal solution of the general problem when the resource constraint is greater than 82 ($k = 13.8$).

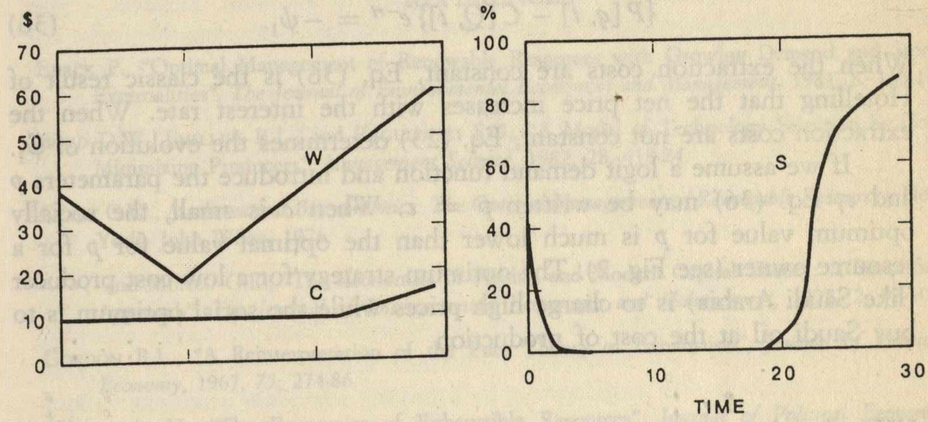

FIGURE 7. The optimal solution of the general problem when the resource constraint is 25 ($k = 26.7$).

case, $K = 26.7$ and the profits are $J = 729$. Most of the production occurs at the end of the period, when the discounted value of the competing price is at a maximum. The rent is large and positive throughout the period.

Equation (24) relates the marginal profits to the first auxiliary variable. If we use a logit demand function, Eq. (24) yields the following bang-bang decision rule: if extraction cost plus rent is less than the price of the competing resource, produce at full capacity; otherwise, stop production.

VII. *The Social Optimum*

Both Hotelling (1931) and Weinstein and Zeckhauser (1975) consider the optimal production of a depletable resource when the objective is to maximize the discounted sum of consumer plus producer surplus. For this case, the objective function is:

$$J = \int_0^T \{ \int_0^q P[y, t]\, dy - C[Q, t]\, q(t)\} e^{-rt}\, dt. \tag{35}$$

The objective function for the social optimum problem is identical to the objective function for the general problem, except that an integral has replaced the product of price and quantity. The solution to the social optimum problem is identical to the solution of the general problem, with the exception that the price elasticity term disappears from Eq. (24):

$$\{P[q, t] - C[Q, t]\} e^{-rt} = -\psi_1. \tag{36}$$

When the extraction costs are constant, Eq. (36) is the classic result of Hotelling that the net price increases with the interest rate. When the extraction costs are not constant, Eq. (23) determines the evolution of ψ_1.

If we assume a logit demand function and introduce the parameters p and c, Eq. (36) may be written $p = c$. When c is small, the socially optimum value for p is much lower than the optimal value for p for a resource owner (see Fig. 2). The optimum strategy for a low cost producer (like Saudi Arabia) is to charge high prices while the social optimum is to buy Saudi oil at the cost of production.

VIII. *Conclusions*

In this paper, we have considered the problem of determining the optimal production path for a depletable natural resource. The classical result of Hotelling is that when the market is in equilibrium, the net price paid to the owner of the resource must increase with the interest rate. We have considered three problems: the basic problem, the general problem, and the social optimum problem. For the basic problem, the market will typically not be in equilibrium, and the classical solution does not yield a production rule. For the general problem and the social optimum problem, the classical solution does work, and we have extended the classical solution to the case of increasing extraction costs.

For the basic problem, we have used the Pontryagin Maximum Principle to find a bang-bang solution for the production rate. The classical solution yields a value for switch points, at which production should be turned on or off.

For all three problems, we have found a differential equation that determines the rent that a resource owner should charge to maximize his profits. The magnitude of the rent depends on its initial value. The proper initial value depends on the total resource and the level of demand.

For the general problem with a logit demand function, the classical solution results in a bang-bang production schedule. Thus, we have found two cases where the optimal production path for depletable natural resources is bang-bang.

REFERENCES

BERCK P., "Optimal Management of Renewable Resources with Growing Demand and Stock Externalities", *The Journal of Environmental Economics and Management*, 1981, 8, 105-17.

BOYD D.W., PHILLIPS R.L., and REGULINSKI S.G., "A Model of Technology Selection by Cost Minimizing Producers", *Management Science*, 1982, 28, 418-24.

CLARK C.W., *Mathematical Bioeconomics: The Optimal Management of Renewable Resources*, New York: John Wiley, 1976.

——, and MUNRO G.R., "The Economics of Fishing and Modern Capital Theory: A Simplified Approach", *The Journal of Environmental Economics and Management*, 1975, *2*, 92-106.

GORDON R.L., "A Reinterpretation of the Pure Theory of Exhaustion", *Journal of Political Economy*, 1967, *75*, 274-86.

HOTELLING H., "The Economics of Exhaustible Resources", *Journal of Political Economy*, 1931, *39*, 137-75.

PINDYCK R.S., "The Optimal Production of an Exhaustible Resource When Price is Exogenous and Stochastic", *Scandinavian Journal of Economics*, 1981, *32*, 277-88.

——, "Uncertainty in the Theory of Renewable Resource Markets", *Review of Economic Studies*, 1984, *51*, 289-303.

PONTRYAGIN L.S. et al., *The Mathematical Theory of Optimal Processes*, New York: Interscience, 1962.

REISTER D.B., "An Analysis of Industrial Demand for Natural Gas", *Energy*, 1983, *8*, 749-55.

SAGAN H., *Introduction to the Calculus of Variations*, New York: McGraw-Hill, 1969.

Solow R.M., "The Economics of Resources or the Resources of Economics", *The American Economic Review*, 1974, *64*, 1-14.

Stiglitz J.E., "Monopoly and the Rate of Extraction of Exhaustible Resources", *The American Economic Review*, 1976, *66*, 655-61.

Weinstein M.C. and Zeckhauser R.J., "The Optimal Consumption of Depletable Natural Resources", *Quarterly Journal of Economics*, 1975, *89*, 371-92.

LA PRODUZIONE BANG-BANG DELLE RISORSE NATURALI ESAURIBILI

Questo articolo riconsidera il problema di determinare il sentiero ottimo di produzione per una risorsa naturale esauribile. Il risultato classico di Hotelling è che il proprietario della risorsa è indifferente se produrre o non produrre quando il prezzo netto della risorsa aumenta con il tasso di interesse. Tuttavia, i forti aumenti e diminuzioni del prezzo del petrolio nell'ultimo decennio illustrano che il prezzo netto può non sempre aumentare con il tasso di interesse. Quando il prezzo netto non aumenta con il tasso di interesse il principio di massimo di Pontryagin può essere usato per estendere il risultato classico a un problema con un programma di produzione bang bang e a problemi con costi di estrazione crescenti.

Rivista Internazionale di Scienze Economiche e Commerciali
Volume 39 (1992), N. 1, 23-48

SPATIAL COMPETITION AND THE ADOPTION OF NEW TECHNOLOGY

by
FABIO MAZZOLA *

1. *Introduction*

Existing paradigms of spatial diffusion offer reasonable interpretations of new technology adoption lags across locations and have been widely applied in empirical studies to explain the shape of the diffusion curve through time and/or space. In existing approaches, however, agents do not really make the choice of adoption and external factors, such as the frequency and the number of contacts in epidemiologic-hierarchic (Hagerstrand-type) models or the critical threshold of the characteristic in the so-called *probit* models, determine the time and the spatial pattern of adoption across space. If we look at the process of diffusion as a sequence of decisions of adoption taken by individual firms inside an industry it is important to consider the adoption choice as directly related to the characteristics of the market structure since the presence of competitors affects considerably the individual choice. In an oligopolistic setting, for instance, we may address the issue of how the firm can modify its locational choice or delay the introduction of a new technology on the basis of a maximization objective which fully takes into account the rivals' optimal choice.

Strategic considerations or market interactions are not the only ones that determine the choice and the time of adoption. Firm characteristics and non market interactions are obviously of paramount importance in such decisions. A strategic approach may offer, nonetheless, essential insights on important forces governing the adoption choice of the firms which may be

* Università di Palermo and University of Pennsylvania.
I thank Professor Masahisa Fujita for useful comments.

further modeled in a way that takes into account interesting extensions motivated by real world phenomena.

When restricted to a discrete number of agents, these considerations call for a game-theoretic (strategic) approach to the problem of spatial diffusion, thus extending, with the inclusion of locational considerations, the approach originated with Reinganum (1981).

A spatial extension of the game theoretic approach to the diffusion of new technology may proceed by first assuming the firms in fixed locations choosing time of adoption and price.

A more ambitious stage of the analysis would consider the adoption choice inside a location-price competition context. Namely, after the adoption of any of the two firms, a location-price game is played with the firms possibly relocating in space. In analyzing this second situation, however, the issue of the existence of equilibrium in a horizontal differentiated product space with asymmetric costs must be explicitly addressed.

This paper shows how the spatial competition approach can be used to analyze the strategic adoption of new technology in a Hotelling-type duopoly. The analysis of spatial price competition (Sec. 2) is followed by the examination of the spatial locational-price competition framework (Sec. 3) in which firms choose endogenously adoption time, location and price in a strategic setting. Problems related with the existence of equilibrium are examined in Sec. 4. Potential extensions and the usefulness of the approach in tackling real world phenomena are discussed in the concluding section.

2. *The Adoption of New Technology in a Spatial Price Competition Context*

Suppose that, in a Hotelling framework, an innovation (cost-reducing or product-improving) is announced for sale and is available at all locations at a particular time. Is it possible to generate a "spatial diffusion", i.e. a sequence of firms' adoptions at different locations and at different dates, purely as a result of strategic interactions among firms? Namely, can a "spatial diffusion" be generated without necessarily considering the differences among firms in terms of size, information structure, learning ability, etc.? An analogous question addressed by Reinganum (1981) in a non spatial context produced as a result that the "diffusion" (adoption at different dates) equilibrium occurs under certain (sufficient) conditions on the benefits (profit gains) and costs related to the innovation.

As a starting point for the analysis of strategic diffusion in a spatial context we can verify whether the Reinganum conditions hold for the Hotelling model in the spatial price competition case.

Assume then the following setup:

A) *Standard Hotelling Assumptions*

a) a continuum of consumers is distributed uniformly at unit density on a bounded line of length 1;

b) there are two producers in the market, supplying the same product at distinct *fixed* locations s_1 and s_2 and charging mill prices p_1 and p_2;

c) each consumer purchases one and only one product per unit time, i.e. there is perfectly inelastic demand;

d) the willingness to pay (to firm *i*) of a representative consumer located at some point x along the line is:

$$w_i = v - p_i - m\,x \qquad [2.1]$$

where v is the valuation of the good and m is the transport cost rate which is assumed to be constant and paid by the consumers. Assume that $v > \max(p_i + m\,x)$ so that w_i is always bigger than zero (i.e., the consumer always buys); each consumer will buy from the producer who charges the lowest delivered price (for which w_i is maximum);

e) each producer faces the same constant marginal cost $\hat{c}$ which is independent of location. There are no fixed costs;

f) (no undercutting condition): the locations of the firms s_1 and s_2 are such that:

$$\{1 + [(s_i - s_j)/3]\}^2 \geq 4/3\,(s_i + 2\,s_j) \qquad i = 1,2;\ j = 1,2 \qquad [2.2]$$

i.e., firms are not too close so that the equilibrium in prices always exists.

B) *Reinganum setup and assumptions*

g) Firms make their choices in a continuous time setting. For $t < T$, producers compete in prices following Nash pure strategies. At time T, a cost reducing innovation is announced for sale and it is made available at any location. If a firm adopts, the marginal cost is reduced to $\underline{c} < \hat{c}$ and stays at $\underline{c}$ forever. Firms must choose an adoption date which cannot be changed after observing the rivals' decision. A new game in prices is played after each adoption. Let's denote the profit allocations (excluding the innovation costs) as π_0 for each firm when no firm has adopted; π_1 and π_2 for the firm leading and, respectively, following in adoption; π_3 for each firm when both firms have adopted.

h) the present value of bringing the innovation in line by time t (add-ing cost of purchase of the innovation and cost of adjustment) is $a(t)$. It is assumed that this function exists and is continuously differentiable over $t \in (0, \infty)$ though purchases of the innovation may take place at selected dates. Specifically, we may assume with Reinganum that:

$$h_1)\ a(t) \geq 0 \quad \text{for any } t \in (0, \infty) \tag{2.3}$$

$$h_2)\ a'(t) < 0;\ \lim a'(t) > 0 \quad \text{as } t \to \infty \tag{2.4}$$

$$h_3)\ a''(t) > r(\pi_1 - \pi_0)\exp(-rt) \quad \text{for any } t \in (0, \infty) \tag{2.5}$$

$$h_4)\ a'(0) < (\pi_2 - \pi_3) \tag{2.6}$$

where r is the interest rate [1]. In order to investigate whether or not these conditions can determine spatial diffusion, let us observe that, on the basis of the adoption decisions of the firms, three different regimes are (potentially) identified. In the first (regime I), none of the two firms has adopted and they both have a marginal cost $\hat{c}$; in the second (regime II), one of the two firms has adopted and has marginal cost $\underset{\smile}{c}$ while the other still keeps a marginal cost $\hat{c}$; in the third (regime III) both firms have adopted and they both have a marginal cost $\underset{\smile}{c}$. Let us compute the profit allocations (excluding the innovation costs) in the three regimes.

Regime I. – If firms are located at s_1 and s_2, the market boundary between the two firms is obtained from the relationship for the marginal consumer, z, namely:

$$p_1 + mX = p_2 + mY \tag{2.7}$$

and the identity:

$$s_2 = s_1 + X + Y \tag{2.8}$$

where X is the distance between s_1 and z and Y between s_2 and z.

After solving for X and Y, the demands for firms 1 and 2 are given by:

$$D_1 = (p_2 - p_1)/2m + (s_1 + s_2)/2 \tag{2.9}$$

[1] Assumptions h_1 and h_2 indicate that the discounted cost of the innovation is positive and decreasing. However, the decrease in cost cannot continue forever. Assumption h_3 indicates that, after a certain time, any further prolongation of the adjustment process will begin to increase costs. Namely, there is an optimal adoption date after which the cost of adoption will increase. In other terms, non adoption is not an alternative. Assumption h_4 implies that the reduction in cost by adopting at time 0 is, in absolute value, bigger than the loss of profit occurring by being a follower. This means that the cost of adoption declines very rapidly and that immediate adoption by any of the firm is ruled out by assumption.

$$D_2 = [(p_1 - p_2)/2m] + 1 - (s_1 + s_2)/2 \qquad [2.10]$$

and the maximization of the profit functions with respect to p_1 and p_2 respectively yields, after some manipulations, the equilibrium prices:

$$p_1^* = \hat{c} + m\,(2 + s_1 + s_2)/3 \qquad [2.11]$$

$$p_2^* = \hat{c} + m\,(4 - s_1 - s_2)/3 \qquad [2.12]$$

and profits:

$$\pi_1\,(\mathrm{I}) = m\,(2 + s_1 + s_2)^2/18 \qquad [2.13]$$

$$\pi_2\,(\mathrm{I}) = m\,(4 - s_1 - s_2)^2/18 \qquad [2.14]$$

Notice that $\pi_1\,(\mathrm{I})$ and $\pi_2\,(\mathrm{I}) > 0$ and that if $s_2 = 1 - s_1$ we are left with $\pi_1\,(\mathrm{I}) = \pi_2\,(\mathrm{I}) = (1/2)\,m$.

Regime II. – The argument runs parallel to the previous analysis as far as the demand for both firms is concerned. In this case, assuming that firm 1 has adopted, maximization of $\pi_1\,(\mathrm{II})$ and $\pi_2\,(\mathrm{II})$ with respect to p_1 and p_2, respectively, yields equilibrium prices:

$$p_1^* = [(\hat{c} + 2\underline{c})/3] + m\,(2 + s_1 + s_2)/3 \qquad [2.15]$$

$$p_2^* = [(\underline{c} + 2\hat{c})/3] + m\,(4 - s_1 - s_2)/3 \qquad [2.16]$$

Notice that:

$$p_2^* - p_1^* = [2m\,(1 - s_1 - s_2)/3] + (\hat{c} - \underline{c})/3 \qquad [2.17]$$

By substituting in D_2 we obtain:

$$D_2 = 1 - [(s_1 + s_2)/2] - [(\hat{c} - \underline{c})/6m] - (1 - s_1 - s_2)/6m \qquad [2.18]$$

Assume, for simplicity, that $s_1 = 1 - s_2$. Then, in order to be $D_2 > 0$ (firm 2 stays in the market) it must be that:

$$1/2 > (\hat{c} - \underline{c})/6m \qquad [2.19]$$

or:

$$(\hat{c} - \underline{c}) < 3m \qquad [2.20]$$

Hence, the difference in cost cannot be too big for both firms to exist in the market.

Profit allocations (excluding innovation costs) at optimal prices are given by:

$$\pi_1\,(\mathrm{II}) = (1/18)\, m\, [(\hat{c} - \underset{\smile}{c}) + m\, (2 + s_1 + s_2)]^2 \qquad [2.21]$$

$$\pi_2\,(\mathrm{II}) = (1/18)\, m\, [(\underset{\smile}{c} - \hat{c}) + m\, (4 - s_1 - s_2)]^2 \qquad [2.22]$$

or:

$$\pi_1\,(\mathrm{II}) = (m/18)\,(2 + s_1 + s_2)^2 + \\ + [(\hat{c} - \underset{\smile}{c})/3]\,\{[(\hat{c} - \underset{\smile}{c})/6m] + (2 + s_1 + s_2)/3\} \qquad [2.23]$$

$$\pi_2\,(\mathrm{II}) = (m/18)\,(4 - s_1 - s_2)^2 + \\ + [(\underset{\smile}{c} - \hat{c})/3]\,\{[(\underset{\smile}{c} - \hat{c})/6m] + (4 - s_1 - s_2)/3\} \qquad [2.24]$$

From [2.21] and [2.22] we notice that π_1 (II) and π_2 (II) > 0 while comparing [2.13] and [2.14] with [2.23] and [2.24] we obtain, for any s_1 and s_2:

$$\pi_1\,(\mathrm{II}) > \pi_1\,(\mathrm{I}) \qquad [2.25]$$

and, for $s_1 = 1 - s_2$:

$$\pi_2\,(\mathrm{I}) > \pi_2\,(\mathrm{II}) \qquad [2.26]$$

The second inequality follows from the fact that the second term in [2.24] is negative if we take into account the restrictions $(\hat{c} - \underset{\smile}{c}) < 3m$ and $s_1, s_2 \in [0, 1]$.

Finally, when $s_1 = (1 - s_2)$, [2.21] and [2.22] reduce to:

$$\pi_1\,(\mathrm{II}) = (1/2m)\,\{[(\hat{c} - \underset{\smile}{c})/3] + m\}^2 \qquad [2.27]$$

$$\pi_2\,(\mathrm{II}) = (1/2m)\,\{[(\underset{\smile}{c} - \hat{c})/3] + m\}^2 \qquad [2.28]$$

Regime III. – In the case in which both firms have adopted, the marginal cost for each firm will be $\underset{\smile}{c}$. The results obtained in regime I apply also to this case since profit allocations are independent of the marginal costs. Hence we may write:

$$\pi_1\,(\mathrm{III}) = \pi_1\,(\mathrm{I}) = m\,(2 + s_1 + s_2)^2/18 \qquad [2.29]$$

$$\pi_2\,(\mathrm{III}) = \pi_2\,(\mathrm{I}) = m\,(4 - s_1 - s_2)^2/18 \qquad [2.30]$$

which, when $s_2 = 1 - s_1$, reduce to: π_1 (III) $= \pi_2$ (III) $= (1/2)/m$. It follows that π_1 (III), π_2 (III) > 0 and that π_1 (II) $> \pi_1$ (III) $= \pi_1$ (I) and π_2 (III) $> \pi_2$ (II).

We can summarize the results as follows. Assuming that firm 1 adopts at time $\hat{t}$ and that firm 2 adopts at time $t^* > \hat{t}$ total payoffs of the two firms will be:

$$\pi_1 = \int_0^{\hat{t}} \{m\,(2 + s_1 + s_2)^2/18\}\,dt + $$
$$+ \int_{\hat{t}}^{t^*} \{(1/18)\,m\,[(\hat{c} - \underset{\smile}{c}) + m\,(2 + s_1 + s_2)]^2\}\,dt + $$
$$+ \int_{t^*}^{\infty} \{m\,(2 + s_1 + s_2)^2/18\}\,dt - a\,(\hat{t}) \qquad [2.31]$$

$$\pi_2 = \int_0^{\hat{t}} \{m\,(4 - s_1 - s_2)^2/18\}\,dt + $$
$$+ \int_{\hat{t}}^{t^*} \{(1/18)\,m\,[(\underset{\smile}{c} - \hat{c}) + m\,(4 - s_1 - s_2)]^2\}\,dt + $$
$$+ \int_{t^*}^{\infty} \{m\,(4 - s_1 - s_2)^2/18\}\,dt - a\,(t^*) \qquad [2.32]$$

Also, profit allocations of the different regimes excluding innovation costs satisfy the following relations:

a) $\pi_i\,(j) > 0 \qquad i = 1, 2;\ j = \text{I, II, III}$ [2.33]

b) $\pi_1\,(\text{II}) > \pi_1\,(\text{III}) = \pi_2\,(\text{III}) = \pi_1\,(\text{I}) = \pi_2\,(\text{I}) > \pi_2\,(\text{II})$ [2.34]

Notice that if firm 1 is the second to adopt, it will receive a gross profit allocation of $\pi_2\,(\text{II})$.

Let us compute:

$$D = [\pi_1\,(\text{II}) - \pi_1\,(\text{I})] - [\pi_1\,(\text{III}) - \pi_2\,(\text{II})] \qquad [2.35]$$

as the difference between the gross profit gain in being the first to adopt and the gross profit gain in being the second to adopt. Assume, for simplicity of calculations, that $s_2 = 1 - s_1$. By using expressions [2.27], [2.13], [2.29] and [2.28] we obtain, after elimination of some terms:

$$D = [(\hat{c} - \underset{\smile}{c})/3] + [(\hat{c} - \underset{\smile}{c})^2/18m] - [(\underset{\smile}{c} - \hat{c})/3] - (\underset{\smile}{c} - \hat{c})^2/18m \qquad [2.36]$$

which reduces to:

$$D = [(\hat{c} - \underset{\smile}{c})^2 + (\underset{\smile}{c} - \hat{c})^2]/18\,m > 0 \qquad [2.37]$$

Hence, the return of being first is positive in the spatial price competition model.

Under assumptions [2.3] – [2.6], profit allocations satisfying conditions [2.33], [2.34] and the requirement that $D > 0$, Reinganum proved that the two firms will adopt with certainty at two different dates, i.e., there is a "diffusion" in the adoption purely as a result of strategic interactions among firms.

More precisely, there are two Nash equilibria in a game in which firms choose, as strategies, the adoption dates. Namely, with probability one half firm 1 will adopt at a time $\hat{t}$ in (T, ∞) and firm 2 at a different time $t^* > \hat{t}$ and with probability one half the result will be reversed.

Although we cannot say a-priori which one of the two equilibria will occur, it is never a Nash equilibrium for the firms to adopt at the same date.

Our results support the conclusions that, with symmetric locations, the Hotelling spatial price competition framework satisfies the Reinganum conditions yielding strategic diffusion equilibrium. A spatial diffusion may then occur as a result of the strategic interaction between the two firms when firms precommit themselves a-priori to an adoption date.

The rationale for precommitment in the adoption context has been questioned in a later paper by Fudenberg and Tirole (1985). They argue that in the Reinganum model precommitment forces the firms to choose an adoption date without possibility of changing it after observing their rivals' actions. Indeed, if we allow firms to react to the rivals' decision, the "diffusion" equilibrium is usually accompanied by the preemption of the firm that adopts earlier. Preemption means that firms adopt sooner than they would were their rival's adoption dates fixed.

The authors show that the diffusion equilibrium is only one of the possible ones. In fact, there is also the possibility of a "late adoption" equilibrium in which both firms adopt simultaneously as late as possible. Usually, the "diffusion" equilibrium is typical of the case in which the leader (first-mover) advantage is very high. The "late adoption" equilibrium, instead, characterizes the markets in which the industry profit is unaffected by the number of firms that have adopted and the follower has a strong incentive to adopt the innovation after the other firm has adopted (the innovation simply transfers the profit from one firm to another one). It may be interesting to investigate whether or not one or the other subcase occurs in the Hotelling framework. It turns out that the answer will depend on the value of the parameters.

Assuming $s_2 = 1 - s_1$, the industry profit under the first regime is equal to:

$$\pi_1 (\mathrm{I}) + \pi_2 (\mathrm{I}) = \pi (\mathrm{I}) = (1/2)\, m + (1/2)\, m = m \qquad [2.38]$$

The same industry profit also characterizes the third regime. Under the second regime, the industry profit will be:

$$\pi_1 (\mathrm{II}) + \pi_2 (\mathrm{II}) = (1/2m) \{[(\hat{c} - \underset{\sim}{c})/3] + m\}^2 + \\ + (1/2m) \{[(\underset{\sim}{c} - \hat{c})/3] + m\}^2 = m + (\hat{c} - \underset{\sim}{c})^2/9m \qquad [2.39]$$

which, under the restriction that $(\hat{c} - \underset{\sim}{c}) < 3m$ yields that $\pi (\mathrm{II}) < 2m$.

Whether or not industry profit is independent on the number of adoptants will eventually depend on the values of the cost difference and of the transportation rate. Given the relationship between m and $(\hat{c} - \underset{\sim}{c})$, the bigger is m, the more likely is the diffusion equilibrium. However, in order to satisfy the price equilibrium condition, the difference in cost cannot be too big with respect to the transportation rate, thus imposing another restriction for the existence of a diffusion equilibrium in the spatial competition context when precommitment is allowed [2].

3. *The Adoption of New Technology in a Spatial Locational-Price Competition Context*

A more ambitious study of "spatial diffusion" should make the locational choice endogenous and address the issue in the spatial price-locational competition context. While keeping the same setup of the previous paragraph, we may assume that firms compete in both prices and locations and that, when any of the two firms adopts, the location-price game is played again until a new equilibrium is reached. By considering the profit allocations originated before and after any adoption by one of the firms (or both) we can investigate whether the assumptions necessary to obtain a diffusion equilibrium are satisfied or not.

Once again, three regimes will, in principle, originate. In the second regime, however, when only one of the two firms has adopted, we face the problem concerning the existence of a location-price equilibrium. A common way to avoid the equilibrium problems in the location-price game is to assume that the game is played sequentially. In the first and third regimes, when marginal costs are equal for both firms, a sequential subgame (perfect)

---

[2] The same analysis performed here for the case of cost-reducing innovation applies entirely to the case in which the innovation is not cost-reducing but quality-improving. For details see TIROLE (1988).

equilibrium when transportation costs are quadratic in distance is given by (Gabszewicz and Thisse, 1986):

$$(p_1^*, s_1^*); (p_2^*, s_2^*) = (t, 0); (t, 1) \qquad [3.1]$$

where t is the transportation rate.

In the second regime, however, the asymmetry between firms will endanger severely the existence of an equilibrium. The firm with lower marginal cost will tend to charge a lower price and locate closer to the other firm in order to capture the whole market. This cannot be an equilibrium since the firm with higher cost can regain positive market share by relocating further away.

The problem has received a formal treatment by Schulz and Stahl (1985). With circular location (product) space and quadratic transportation costs they show that, in the three firm case, no sequential location-price equilibrium exists when firms have different marginal costs. They also claim that the results extend to a larger number of firms and to the case of linear bounded location space. The general reason for this result is that the firms' payoffs depend on the (relative) distance to the other firms in the location space rather than on the (absolute) location of the firm. In the two firm case, the payoff of the firm with low cost technology is no longer quasi-concave for the reasons outlined before (the firm that has adopted has an incentive to locate very close to the other firm).

In what follows the case of quadratic transportation costs in a linear market will be briefly analyzed for a spatial duopoly in order to verify these hypotheses. To show the result assume that, in the Hotelling context outlined in section 2, the transport cost function is quadratic, i.e., $m(s_i', s_i'') = m(s_i' - s_i'')^2$ for any $s_i', s_i'' \in S \in [0, 1]$.

The utility function of a representative consumer located at x can be represented by:

$$U_i(q, s_i) = q + v - p_i - m(s_i - x)^2 \qquad [3.2]$$

where q is the numeraire commodity and the transportation cost is quadratic in distance from firm's location s_i.

The marginal consumer equation yields the market boundary from which we derive the demand functions:

$$D_1 = [(p_2 - p_1) + m(s_2^2 - s_1^2)]/2m(s_2 - s_1) \qquad [3.3]$$

$$D_2 = [(p_1 - p_2) + m(s_2 - s_1)(2 - s_1 - s_2)]/2m(s_2 - s_1) \qquad [3.4]$$

Let $s_1 \leq s_2$ and firm 1 be the first adoptant. Then, the maximization of the profit functions yields, after some algebra, the equilibrium prices:

$$p_1^* = [(2\underline{c} + \hat{c})/3] + m\,(s_2 - s_1)\,(2 + s_1 + s_2)/3 \qquad [3.5]$$

$$p_2^* = [(\underline{c} + 2\hat{c})/3] + m\,(s_2 - s_1)\,(4 - s_1 - s_2)/3 \qquad [3.6]$$

and the equilibrium profits as functions of locations:

$$\pi_1^*\,(s_1, s_2) = \{[(\hat{c} - \underline{c})/3] + m\,(s_2 - s_1)\,(2 + s_1 + s_2)/3\} \times \times \{\,[(\hat{c} - \underline{c})/6m\,(s_2 - s_1)] + (2 + s_1 + s_2)/6\} \qquad [3.7]$$

$$\pi_2^*\,(s_1, s_2) = \{[(\underline{c} - \hat{c})/3] + m\,(s_2 - s_1)\,(4 - s_1 - s_2)/3\} \times \times \{[(\underline{c} - \hat{c})/6m\,(s_2 - s_1)] + (4 - s_1 - s_2)/6\} \qquad [3.8]$$

Profits allocations can also be written, after some manipulations:

$$\pi_1^*\,(s_1, s_2) = [m\,(s_2 - s_1)\,(2 + s_1 + s_2)^2/18] + + [(\hat{c} - \underline{c})/18m\,(s_2 - s_1)] + [2\,(\hat{c} - \underline{c})\,(2 + s_1 + s_2)/18] \qquad [3.9]$$

$$\pi_2^*\,(s_1, s_2) = [m\,(s_2 - s_1)\,(4 - s_1 - s_2)^2/18] + + [(\underline{c} - \hat{c})/18m\,(s_2 - s_1)] + [2\,(\underline{c} - \hat{c})\,(4 - s_1 - s_2)/18] \qquad [3.10]$$

If the cost difference is zero, straightforward maximization of the first and third terms of both equations with respect to the location choice yields $s_1^* = 0$ and $s_2^* = 1$. On the other hand, when firms have different costs, the second term in [3.7] is maximized when $s_1 = s_2 - \epsilon$. Hence, firm 1 has a strong incentive to go very close to firm 2 in order to drive the second firm out of the market. This is precisely the fact pointed out by Schulz and Stahl (the payoffs of the firms depend on the relative distance and not on the location proper). The degeneracy of the payoff with respect to location choice is the reason for non existence of the equilibrium in the first stage of the game (locational equilibrium). In our case, firm 2 will try to move further from firm 1 to restore positive market share [3].

Given this non existence result, some modifications of the standard Hotelling framework must be introduced in order to ensure equilibrium in the second regime when firms have different costs. In the following section we analyze the potential role of slight variations in the assumptions concerning utility functions, consumers' density distribution and price setting behav-

[3] The non-existence problem also applies to the case in which the asymmetry of the firm is due to the adoption of a quality-improving innovation by one of them. The argument runs parallel to the one above presented with δv (the quality improvement) now taking the place of $\delta c = \hat{c} - \underline{c}$.

ior. Later, in the concluding section, we discuss the more fruitful possibility of enlarging the general framework of the analysis bypassing the rigidity of the traditional Hotelling context.

4. *Adoption, Cost Asymmetry and Location-Price Equilibrium*

4.1. *Consumers' Heterogeneity*. – In their reformulation of the Hotelling problem, De Palma, Ginsburgh, Papageorgiou and Thisse (1985), show how consumers' heterogeneity can help to restore both equilibrium and central agglomeration in the location-price competition context. Their approach can be summarized as follows. Each of the two duopolists cannot predict with certainty the behavior of a particular consumer. Rather, the utility of a consumer at s purchasing from firm i can be expressed as:

$$u_{is} = v - p_i - t(s, s_i) + \epsilon_i \qquad i = 1, 2 \qquad [4.1]$$

where ϵ_i is Weibull distributed with mean 0 and standard deviation σ. Firms can evaluate only the probability P_{is} that a consumer at s will purchase from firm i as:

$$P_{is} = \text{Prob}\,(u_{is} = \max_{j=1,2} u_{js}) \qquad [4.2]$$

Since $\epsilon_i \sim$ Weibull $(0, \sigma)$, we may express the probability in a logit fashion:

$$P_{is} = \exp\{-[p_i + t(s, s_i)]/\sigma\}/\sum_{j=1}^{2} \exp\{-[p_i + t(s, s_i)]/\sigma\} \qquad [4.3]$$

Suppose a bounded line of unit length. For the first firm, assuming linear transportation costs, [4.3] reduces to:

$$P_1 = 1/\{1 + [\exp(-(p_2 + t\,|\,s - s_2\,|]/\sigma)/ \\ /\exp(-[p_1 + t\,|\,s - s_1\,|]/\sigma)\} \qquad [4.4]$$

Let us assume, with no loss of generality, that $s_1 \leq s_2$. The value of [4.4] differs according to the consumers' location in [0, 1]. For $0 \leq s < s_1$, it is possible to show, algebraically, that:

$$P_{1A} = 1/[1 + \exp(A)] \qquad [4.5]$$

where $A = (1/\sigma)\,[(p_1 - p_2) + t(s_1 - s_2)]$.

For $s_2 < s \leq 1$ we are left with:

$$P_{1C} = 1/[1 + \exp(C)] \qquad [4.6]$$

where $C = (1/\sigma)\,[(p_1 - p_2) + t\,(s_2 - s_1)]$.

Finally, for $s_1 \leq s \leq s_2$, we have:

$$P_{1B} = 1/[1 + \exp(B)] \quad [4.7]$$

where $B = (1/\sigma)\,[(p_1 - p_2) + t\,(2s - s_1 - s_2)]$.

Total demand for firm 1 will be given by:

$$D_1 = \int_0^{s1} P_{1A}(s)\,ds + \int_{s1}^{s2} P_{1B}(s)\,ds + \int_{s2}^{1} P_{1C}(s)\,ds \quad [4.8]$$

Calculation of integrals yields, after some manipulations:

$$D_1 = \{s_1/[1 + \exp(A)]\} + \{(1 - s_2)/[1 + \exp(C)]\}$$
$$- (\sigma/2t)\log\{[1 + \exp(-A)]/[1 + \exp(-C)]\} \quad [4.9]$$

A symmetric argument yields the demand for firm 2:

$$D_2 = \{s_2/[1 + \exp(A^*)]\} + \{(1 - s_1)/[1 + \exp(C^*)]\}$$
$$- (\sigma/2t)\log\{[1 + \exp(-A^*)]/[1 + \exp(-C^*)]\} \quad [4.10]$$

where:

$$A^* = (1/\sigma)\,[(p_2 - p_1) + t\,(s_2 - s_1)] \quad [4.11]$$

$$B^* = (1/\sigma)\,[(p_2 - p_1) + t\,(s_1 + s_2 - 2s)] \quad [4.12]$$

$$C^* = (1/\sigma)\,[(p_2 - p_1) + t\,(s_2 - s_1)] \quad [4.13]$$

Since demand is continuous in both prices and locations, profit allocations will also be continuous in the variables [4].

Maximization of the profit allocations when firms have different costs can then be performed in order to determine the equilibrium for the second regime. In their article, De Palma, Ginsburgh, Papageorgiou and Thisse (1985) show that for high values of σ (in our context bigger than t) the

[4] Notice that $\sigma = 0$ implies that $\pi_1 = \pi_2 = 0$. Also, as σ (the degree of consumers' heterogeneity) increases, profits of both firms will start to increase. If $\sigma \rightarrow \infty$ we are left with: $D_1 = [1 + (s_1 - s_2)]/2$ and: $D_2 = [1 + (s_2 - s_1)]/2$ which are not continuous in locations for $s_1 = s_2$. Hence, we must generally assume that $\sigma > 0$ and finite.

(sequential) location price game generates an equilibrium (one of the possible) with firms charging $p_1^* = p_2^* = 2\sigma$ and located at $s_1 = s_2 = 1/2$. The reason for this result is that, when heterogeneity is high, the demand to each firm becomes increasingly less dependent on their rivals' location. Hence firms do not need to go further to soften competition. The conclusion could be, in principle, extended to the case in which firms have asymmetric costs although, in this context, it seems unlikely to obtain equal prices at equilibrium. For $\sigma < t$ De Palma, Ginsburgh, Papageorgiou and Thisse (1985) obtain inconclusive results about the tendency of firms to agglomerate. In a slight different framework [5] Braid (1988) obtains numerical results showing that, even with moderate values of σ, the two firm case yields agglomeration at the center.

Since firms tend to agglomerate at equilibrium, the framework should not allow for strategic adoption at different times and different locations, at least in the duopoly context.

Explicit solutions for low heterogeneity settings are needed to confirm this conjecture.

4.2. *Consistent Price Conjectures.* – Schulz and Stahl (1985) have speculated that problems of non existence of locational equilibrium in the spatial competition context may vanish when the distribution of consumers is not uniform. An extreme case of non-uniformity may be given by a market in which consumers are located at mass points (for instance, at the extremities of a bounded line of length 1).

With atomic markets, however, the demand function to a firm is highly discontinuous and the solution of the price game cannot be obtained (Gabszewicz and Thisse, 1986). For this reason, an alternative concept of price equilibrium must be considered in order to solve the second stage of the game.

Eaton and Kierzkowki (1984) have adapted the concept of consistent conjectures equilibrium to a spatial duopoly following a previous work by Bresnahan (1981). Loosely, a consistent conjectures equilibrium is one in which a firm, instead of having Cournot conjectures, takes into account the fact that a variation of prices makes it viable for the rival to consider the possibility of undercutting its price. The result is that, at equilibrium, each firm will charge the biggest price it can without provoking a price cut by the other firm.

[5] Braid considers a spatial location competition context in which prices are exogenous. Also, he allows consumers to be located at five different points along the line and not along a continuum.

Consider a duopoly with firms serving consumers located at 0 and 1 (extremities) of the market area, with dimensionality n_1 and n_2, respectively. Assume, with no loss of generality, that $n_1 + n_2 = 1$ and that $n_1 \geq n_2$. The two firms must choose locations on the line (first stage) and price (second stage). Let denote the firms' locations by s_1 and s_2, respectively [s_1, $s_2 \in [0, 1]$]. Allow firms to have different marginal costs, c_1 and c_2, respectively. As usual, suppose that consumers buy from the firm that offers the lowest price (inclusive of transportation cost – at a rate t – from the consumers' to the firm's location). Suppose for the moment that these prices are such that firm 1 sells to the n_1 consumers located at 0 and firm 2 to the n_2 consumers located at 1. Following Eaton and Kierzkowski (1984) suppose that firm 1 conjectures that a departure from the equilibrium price, p_1^*, will generate a response from firm 2 of the type:

$$\begin{aligned} p_2(p_1) &= p_2^* && \text{if } p_1 \leq p_1^* \\ &= p_1 - t(s_2 - s_1) && \text{if } p_1 > p_1^* \end{aligned} \quad [4.14]$$

namely, an increase from the equilibrium price will push the rival to undercut the new price. Given this conjecture, the optimal value of p_1 will be:

$$\begin{aligned} p_1 &= p_1^* && \text{if } (p_1^* - c_1)\, n_1 \geq [p_2^* - t(s_2 - s_1)] - c_1 \\ &= p_2^* - t(s_2 - s_1) && \text{if } (p_1^* - c_1)\, n_1 < [p_2^* - t(s_2 - s_1)] - c_1 \end{aligned} \quad [4.15]$$

If firm 2 conjectures the response of firm 1, it behaves accordingly to its optimal response. More precisely, it fixes p_2^* in a way that firm 1 will be forced to stay and not to undercut. Hence we may write:

$$(p_1^* - c_1)\, n_1 = p_2^* - t(s_2 - s_1) - c_1 \quad [4.16]$$

from which:

$$p_2^* = p_1^* n_1 + c_1 n_2 + t(s_2 - s_1) \quad [4.17]$$

Symmetrically, for firm 1 we have:

$$p_1^* = p_2^* n_2 + c_2 n_1 + t(s_2 - s_1) \quad [4.18]$$

Solving the system defined by [4.17] and [4.18] we obtain the consistent conjectures equilibrium prices:

$$p_1^0 = [n_2^2 c_1 + n_1 c_2 + (1 + n_2)\, t(s_2 - s_1)]/(1 - n_1 n_2) \quad [4.19]$$

$$p_2^0 = [n_1^2 c_2 + n_2 c_1 + (1 + n_1)\, t(s_2 - s_1)]/(1 - n_1 n_2) \quad [4.20]$$

from which we obtain (second stage) equilibrium profits:

$$\pi_1^* = (p_1^0 - c_1)\, n_1 \qquad [4.21]$$

$$\pi_2^* = (p_2^0 - c_2)\, n_2 \qquad [4.22]$$

Notice that, without considering cost differences, if $n_1 > n_2$ we have that $p_2^* > p_1^*$ and $\pi_1^* > \pi_2^*$. Hence, the firm with the higher market will obtain a higher profit [6].

By maximizing [4.21] and [4.22] with respect to locations we obtain the location-price equilibrium $(p_1^*, s_1^*; p_2^* s_2^*) = (0, p_1^0; 1, p_2^0)$ which allows for asymmetry in (marginal) costs between the two firms and can be used for the definition of the profit allocations in regime II. Indeed, some caveats must be kept in mind. First, we have assumed that firm 1, which has the cost advantage, will obtain the market with higher number of consumers. This is a plausible assumption. Firm 2 will then be forced not to locate at 0 in order to obtain a positive market share. Since equilibrium price is directly related to the distance between the two firms, firm 2 has an incentive to locate further away from firm 1 in order to increase its profit.

Suppose, however, that no firm has a cost advantage, i.e. $c_1 = c_2 = c$. Both firms will then be tempted to locate at the market with higher number of consumers (if they cannot sell to both markets) or in the middle of the line (in case they can). Hence, they will be located at the same point. With equal costs and $s_1 = s_2$:

$$\hat{p}_2 = \hat{p}_1 = [(n_2^2 + n_1)\, c]/(1 - n_1 n_2) = c \qquad [4.23]$$

Hence, firms will obtain zero profits. This cannot be an equilibrium since firms can obtain positive market shares by relocating further away. The couples $(p_1^*, s_1^*; p_2^*, s_2^*) = (0, \hat{p}_1; 0, \hat{p}_2)$ or $(p_1^*, s_1^*; p_2^*, s_2^*) = [(1/2), \hat{p}_1; (1/2), \hat{p}_2]$ cannot be a location-price equilibrium.

Paradoxically, symmetry in costs may cause non existence of the location-price equilibrium. This non-existence problem could be solved, in principle, by allowing a different setup. A first possibility is to allow a continuous

---

[6] Both profits are, in any case, positive at equilibrium. Price undercutting is prevented in this type of equilibrium. However, we must require that if $c_1 < c_2$, $p_2 > c_2$ otherwise firm 1 will drive firm 2 out of the market by pricing just below c_2. Imposing $p_2^* > c_2$ is equivalent to impose that:

$$n_1^2 c_2 + n_2 c_1 + (1 + n_1)\, t\, (s_2 - s_1) > c_2 (1 - n_1 n_2)$$

which, recalling that $n_2 = 1 - n_1$, yields: $t\,(s_2 - s_1) > (c_2 - c_1)[n_2/(1 + n_1)]$.

distribution of consumers in the market while keeping the assumption of consistent conjecture in the second stage of the game. By this way, the particular attractiveness of some locations could be avoided. In this case, n_1 and n_2 would be substituted by the market areas for firm 1 and firm 2 obtained through the usual marginal condition (see Section 2). We found, however, that for any value of the cost difference between c_1 and c_2, there exists no unique consistent conjectures equilibrium couple of prices [7]. Therefore, the subgame perfect equilibrium concept cannot be applied because we do not have a unique solution for the second stage of the game.

Alternatively, as it is done in Eaton and Kierzkowski (1984), or in Prescott and Visscher (1977), firms may be thought to choose location sequentially. The issue of entry deterrence must then be addressed explicitly. Eaton and Kierzkowski show that for some parameter values (concerning the relation between number of consumers and fixed costs [8]), the incumbent does not find profitable to deter entry. Under particular cases, the first entrant will find profitable to locate where the number of consumers is bigger and the second entrant at the other side of the market. The equations [4.19] and [4.20] may then represent, for any value of c_1 and c_2, the solution to the third stage of a game in which the first two stages are given by the successive location choices of the two firms.

Suppose that firm 1 is, with no loss of generality, the first firm to enter the market under any regime and the first firm to adopt a new technology which brings a decrease in the marginal production cost from c_2 to c_1. With profits given by [4.21] and [4.22] and optimal location choices given by $s_1^* = 0$ and $s_2^* = 1$, profit allocations for firms 1 and 2 under the different regimes are given by:

$$\pi_1(\mathrm{I}) = \{[((n_2^2 + n_1)c_2 + (1 + n)t)/(1 - n_1 n_2)] - c_2\} n_1 \quad [4.24]$$

$$\pi_2(\mathrm{I}) = \{[((n_1^2 + n_2)c_2 + (1 + n_1)t)/(1 - n_1 n_2)] - c_2\} n_2 \quad [4.25]$$

$$\pi_1(\mathrm{II}) = \{[n_2^2 c_1 + n_1 c_2 + (1 + n_2)t]/[1 - n_1 n_2] - c_1\} n_1 \quad [4.26]$$

$$\pi_2(\mathrm{II}) = \{[n_1^2 c_2 + n_2 c_1 + (1 + n_1)t]/[1 - n_1 n_2] - c_2\} n_2 \quad [4.27]$$

---

[7] The algebraic demonstration of this result is involved. By applying the definition of consistent conjecture to the case of continuous distribution of consumers, we obtain the following value for $p_1(p_2)$: $p_1(p_2) = p_2 + ts_2 + [ts_1(c_2 - p_2 - 2t)]/(c_2 - p_2 + 2t)$ which can then be substituted in the equation for p_2 to get a third degree equation in p_2. This equation has a negative discriminant, i.e., three different real roots.

[8] In a model of simultaneous location choice, fixed costs are usually assumed to be equal to zero. When entry is allowed, however, fixed costs cannot be equal to zero in order to have an equilibrium with a finite number of agents.

$$\pi_1\,(\mathrm{III}) = \{[((n_2^2 + n_1)\,c_1 + (1 + n_2)\,t)/(1 - n_1 n_2)] - c_1\}\,n_1 \quad [4.28]$$

$$\pi_2\,(\mathrm{III}) = \{[((n_1^2 + n_2)\,c_1 + (1 + n_1)\,t)/(1 - n_1 n_2)] - c_1\}\,n_2 \quad [4.29]$$

Notice, also, that if firm 1 is follower in adoption and not in entry, it will have a payoff of:

$$\pi_1^F\,(\mathrm{II}) = \{[n_2^2 c_2 + n_1 c_1 + (1 + n_2)\,t]/(1 - n_1 n_2) - c_2\}\,n_1 \quad [4.30]$$

All profit allocations are positive as long as equilibrium prices are bigger than marginal costs. By inspection of [4.24]-[4.30] we obtain the following relationships:

$$\pi_i\,(\mathrm{II}) > \pi_i\,(\mathrm{I}) = \pi_i\,(\mathrm{III}) > \pi_i^F\,(\mathrm{II}) \qquad i = 1,\ 2 \quad [4.31]$$

and:

$$D = [\pi_i\,(\mathrm{II}) - \pi_i\,(\mathrm{I})] - [\pi_i\,(\mathrm{III}) - \pi_i^F\,(\mathrm{II})] = 0 \qquad i = 1,\ 2 \quad [4.32]$$

Hence, the Reinganum assumption requiring a big advantage for being first in adoption as a condition for strategic diffusion is not satisfied.

If firm 1 is follower in both adoption and entry its payoff in regime II will be $\pi_2\,(\mathrm{II})$. Computation of the industry profit for symmetric costs regimes yields the following relationship:

$$\pi_1\,(\mathrm{III}) + \pi_2\,(\mathrm{III}) = \pi\,(\mathrm{III}) = \pi\,(\mathrm{I}) = \pi_1\,(\mathrm{I}) + \pi_2\,(\mathrm{I}) \quad [4.33]$$

We may also notice that:

$$\pi_1\,(\mathrm{II}) - \pi_1\,(\mathrm{I}) = n_1\,(c_2 - c_1) \quad [4.34]$$

$$\pi_2\,(\mathrm{I}) - \pi_2\,(\mathrm{II}) = n_2\,(c_2 - c_1) \quad [4.35]$$

Therefore, if $n_1 = n_2 = 1/2$:

$$\pi_1\,(\mathrm{II}) - \pi_1\,(\mathrm{I}) = \pi_2\,(\mathrm{I}) - \pi_2\,(\mathrm{II}) \quad [4.36]$$

or:

$$\pi_1\,(\mathrm{II}) + \pi_2\,(\mathrm{II}) = \pi\,(\mathrm{II}) = \pi\,(\mathrm{I}) = \pi_1\,(\mathrm{I}) + \pi_2\,(\mathrm{I}) \quad [4.37]$$

Namely, in the case of equal market size, the verification of the Fudenberg and Tirole assumptions for the occurrence of the "diffusion" equilibrium is not satisfied no matter what is the value of the cost difference between the two firms. If $n_1 > n_2$, instead, the bigger is the difference $(n_1 - n_2) > 0$, the bigger will be the industry profit under the second regime and the more likely will be the "diffusion" equilibrium. By subtracting equation [4.35] from [4.36], we obtain, in fact:

$$\pi\,(\mathrm{II}) - \pi\,(\mathrm{I}) = (c_2 - c_1)\,(n_1 - n_2) \qquad [4.38]$$

The consistent conjecture approach coupled with sequential location choice and atomic markets of consumers may then lead to a "spatial diffusion" in the sense previously explained when both price and location choices are made endogenous. Also, conditions on the degree of cost saving due to adoption that generates diffusion are not so crucial as they are in the fixed location context. However, both the weak theoretical foundation of the concept of consistent conjecture (see, in particular, Tirole, 1988) and the limiting assumption of sequential location choice call for the examination of alternative frameworks of analysis.

4.3. *Discriminatory Pricing.* – Discriminatory pricing has been invoked by Gabszewicz and Thisse (1986) as another possible way to escape solution problems in the spatial locational-price competition context. The following setup adapts the general framework developed in Hurter and Lederer (1985) and Lederer and Hurter (1986) to the Hotelling setting of the previous sections. Assume that two firms are located at s_1 and s_2 on a bounded line of unit length. Transportation cost is a function of consumer location x and of firm location and is assumed, with no loss of generality, to be quadratic in distance, i.e. $t_i\,(s_i, x) = t_i\,(s_i - x)^2$ for $i = 1, 2$.

Firms play a sequential location-price game. The second stage of the game is to set a price policy, namely a function $p_1\,(s_1, s_2, x)$ that states that firm 1 will offer a unit of the good to x at p_1 when it is located at s_1 and firm 2 is located at s_2. Allow firms to have different marginal costs. Profit allocations for each of the two duopolists are then:

$$\pi_i\,(s_i, s_j, p_i, p_j) = \int_{p_i < p_j} [p_i\,(s_i, s_j, x) - t_i\,(s_i - x)^2 - c_i]\,dx +$$

$$+ \,(1/2) \int_{p_i = p_j} [p_i\,(s_i, s_j, x) - t_i\,(s_i - x)^2 - c_i]\,dx$$

$$i, j = 1, 2. \qquad [4.39]$$

where:

$$(p_i \leqslant p_j) = \{x \mid p_i\,(s_i, s_j, x) \leqslant p_j\,(s_i, s_j, x)\} \qquad [4.40]$$

We have assumed that firms share the area in which their prices are equal. We solve first the second stage of the game. For any location s_1 and s_2 and price policy $p_2(p_1, s_2, x)$, an optimal price policy for firm 1 is:

$$p_1(s_1, s_2, x) = \max[p_2(s_1, s_2, x) - \epsilon;\ t_1(s_1 - x)^2 + c_1] \quad [4.41]$$

Knowing [4.41], firm 2 will select a price policy such that: a) firm 1 is forced to price at total (production plus transportation) marginal cost and b) firm 2 may undercut it. So we may write:

$$p_2(s_1, s_2, x) = \max[t_1(s_1 - x)^2 + c_1 - \epsilon';\ t_2(s_2 - x)^2 + c_2] \quad [4.42]$$

Knowing [4.42], firm 1 will select an analogous price policy such that:

$$p_1(s_1, s_2, x) = \max[t_2(s_2 - x)^2 + c_2 - \epsilon;\ t_1(s_1 - x)^2 + c_1] \quad [4.43]$$

Optimal prices will then be given by:

$$p_1^*(s_1, s_2, x) = \lim_{\epsilon \to 0} p_1(s_1, s_2, x) = \\ = \max[t_2(s_2 - x)^2 + c_2;\ t_1(s_1 - x)^2 + c_1] \quad [4.44]$$

$$p_2^*(s_1, s_2, x) = \lim_{\epsilon \to 0} p_2(s_1, s_2, x) = \\ = \max[t_1(s_1 - x)^2 + c_1;\ t_2(s_2 - x)^2 + c_2] \quad [4.45]$$

If:

$$t_1(s_1 - x)^2 + c_1 < t_2(s_2 - x)^2 + c_2 \quad [4.46]$$

for some x internal to the interval [0, 1], firm 1 will have a cost advantage on firm 2. According to [4.44] and [4.45], both firms will charge a price of $p_1^* = p_2^* = t_2(s_2 - x)^2 + c_2$. However, firm 1 has the possibility of charging a price slightly lower and obtain all the market of consumers located at x. The opposite will happen when:

$$t_1(s_1 - x)^2 + c_1 > t_2(s_2 - x)^2 + c_2 \quad [4.47]$$

If $t_1 = t_2$, the area for which $t(s_1 - x)^2 + c_1 = t(s_2 - x)^2 + c_2$ will consist of one point only. In any case, according to [4.39], both firms will earn zero profits in this area. When transportation rates are equal, the profit functions under equilibrium price schedules will be equal to:

$$\pi_1^*(s_1, s_2, p_1^*, p_2^*) = \int_0^{\hat{x}} [t(s_2 - x)^2 + c_2 - t(s_1 - x)^2 - c_1]\, dx \quad [4.48]$$

$$\pi_2^* (s_1, s_2, p_1^*, p_2^*) = \int_{\hat{x}}^{1} [t(s_1 - x)^2 + c_1 - t(s_2 - x)^2 - c_2]\, dx \qquad [4.49]$$

where [9]:

$$\begin{aligned}(0 - \hat{x}) &= \{x \mid t(s_2 - x)^2 + c_2 > t(s_1 - x)^2 + c_1\} \\ (\hat{x} - 1) &= \{x \mid t(s_2 - x)^2 + c_2 < t(s_1 - x)^2 + c_1\}\end{aligned} \qquad [4.50]$$

Profit allocation for each firm may also be written from [4.39]:

$$\pi_1^* (s_i, s_j, p_i^*, p_j^*) = \int_0^1 [t(s_j - x)^2 + c_j]\, dx -$$

$$- \int_0^1 \min [t(s_i - x)^2 + c_i;\ t(s_j - x)^2 + c_j]\, dx \qquad i, j = 1, 2 \qquad [4.51]$$

Maximization of [4.51] with respect to s_i yields optimal locations. Hurter and Lederer observe that a locational equilibrium is assured by minimization of the second term of the equation with respect to the location choice since we are in presence of a continuous function defined over a compact space.

In order to obtain explicit solutions we discuss, in turn, specific cases. Notice that when $s_1 = s_2$ and $c_1 = c_2$ it follows that: $\pi_1 = \pi_2 = 0$. Hence, symmetric firms at equilibrium cannot be located at the same point. Indeed, by relocating, firms can regain the total (transportation plus production) cost advantage. Suppose that $s_1 \neq s_2$ but $c_1 = c_2$. Then, $\hat{x} = (s_1 + s_2)/2$. Equilibrium profits in this case will be given by:

$$\pi_1^* = \int_0^{(s_1 + s_2)/2} t[(s_2 - x)^2 - (s_1 - x)^2]\, dx \qquad [4.52]$$

$$\pi_2^* = \int_{(s_1 + s_2)/2}^{1} t[(s_1 - x)^2 - (s_2 - x)^2]\, dx \qquad [4.53]$$

[9] It is assumed that at $\hat{x}$ the following holds: $t(s_1 - \hat{x})^2 + c_1 = t(s_2 - \hat{x})^2 + c_2$, from which: $\hat{x} = [(c_2 - c_1)/2t(s_2 - s_1)] + [(s_1 + s_2)/2]$.

Maximization of [4.52] and [4.53] with respect to location choices yields the first order conditions:

$$t\,[(s_2 - s_1)/2]^2 - t\,(s_1)^2 = 0 \qquad [4.54]$$

$$t\,(s_2 - 1)^2 - t\,[(s_2 - s_1)/2]^2 = 0 \qquad [4.55]$$

The system [4.54]-[4.55] is solved for $s_1^* = 1/4$ and $s_2^* = 3/4$. This solution is unique for $s_1, s_2 \in [0, 1]$. Therefore we may conclude that:

$$\pi_1\,(\mathrm{I}) = \pi_1\,(\mathrm{III}) = \int_0^{1/2} t\{[(3/4) - x]^2 - [(1/4) - x]^2\}\,dx \qquad [4.56]$$

$$\pi_2\,(\mathrm{I}) = \pi_2\,(\mathrm{III}) = \int_{1/2}^{1} t\{[(1/4) - x]^2 - [(3/4) - x]^2\}\,dx \qquad [4.57]$$

Notice that $\pi_1\,(\mathrm{I}) = \pi_1\,(\mathrm{III}) > 0$ and $\pi_2\,(\mathrm{I}) = \pi_2\,(\mathrm{III}) > 0$. In order to determine profit allocations in the asymmetric cost case (assume $c_2 > c_1$), we must maximize [4.48] and [4.49] with respect to the location choice. Namely:

$$\max_{s_1} \pi_1 = \max_{s_1} \int_0^{[(c_2-c_1)/2t\,(s_2-s_1)] + [(s_1+s_2)/2]} \{t\,[(s_2 - x)^2 - (s_1 - x)^2] + (c_2 - c_1)\}\,dx \qquad [4.58]$$

$$\max_{s_2} \pi_2 = \max_{s_2} \int_{[(c_2-c_1)/2t\,(s_2-s_1)] + [(s_1+s_2)/2]}^{1} \{t\,[(s_1 - x)^2 - (s_2 - x)^2] + (c_1 - c_2)\}\,dx \qquad [4.59]$$

After some manipulations, first order conditions reduce to:

$$t\,[((s_1 - s_2)/2) - (c_2 - c_1)/2t\,(s_2 - s_1)]^2 - t\,s_1^2 = 0 \qquad [4.60]$$

$$t\,(s_2 - 1)^2 - t\,[((s_2 - s_1)/2) - (c_2 - c_1)/2t\,(s_2 - s_1)]^2 = 0 \qquad [4.61]$$

which can be solved to determine the equilibrium locations in the asymmetric cost case. By introducing the solutions into [4.48] and [4.49] we can verify whether or not the Reinganum and the Fudenberg and Tirole conditions for the existence of a diffusion equilibrium (with and without precommitment) are satisfied.

Nothing seems to prevent a diffusion equilibrium from being theoretically possible. As usual, the existence of such equilibrium should depend on the importance of the cost differential. It must be observed, however, that the addition of equations [4.60] and [4.61] yields, after elimination of some terms:

$$(c_2 - c_1) + t\,(s_2 - 1)^2 - ts_1^2 = 0 \qquad [4.62]$$

which implies that:

$$(c_2 - c_1) = ts_1^2 - t\,(s_2 - 1)^2 \qquad [4.63]$$

The cost differential [4.63] is maximized when $s_2 = 1$ and $s_1 = 1$. Hence, in order to satisfy the locational equilibrium, we must have:

$$(c_2 - c_1) \leq t \qquad [4.64]$$

Restriction [4.64] reduces the possibility of a diffusion equilibrium. We conclude that, with price discrimination, a strategic spatial diffusion with endogenous location choice is theoretically possible although under very restrictive conditions [10].

5. *Extensions of the Approach and Concluding Remarks*

The main results of the previous analysis may be summarized as follows. The framework developed by Reinganum (1981) and refined by Fudenberg and Tirole (1985) may be adapted to the spatial competition case to show the possibility of a "spatial diffusion" in the adoption of a new technology generated by the strategic behavior of the firms. The occurrence of a "diffusion" equilibrium is positively related to the cost reduction generated by the adoption. With exogenous location, however, cost reduction cannot be too big with respect to the value of the transportation rate in order to satisfy the equilibrium in the price stage of the game. If the locational choice is made endogenous, the framework is of limited applicability because of the problems related with the existence of a location-price equilibrium when firms have asymmetric costs. If firms can price discriminate, existence can be restored and a "spatial diffusion" equilibrium is theoretically possible. However, stronger restrictions apply to the degree

[10] It may be recalled that the necessary condition for a diffusion equilibrium with both firms active in the case of linear transportation cost function was that $(c_2 - c_1) < 3t$. When transportation costs are quadratic it can be shown that this necessary condition is $(c_2 - c_1) < 4t$.

of cost asymmetry in order to satisfy the locational stage of the game. This reduces the range of values of parameters for which a "diffusion" equilibrium will occur.

Formal proofs of some of the results considered are needed in order to better justify these arguments. At a first view, they are not, indeed, counter-intuitive. If firms are allowed to choose location, there is a reduced gain in being a follower in adoption even when the cost of adoption reduces through time since the leader can take advantage of a small difference in costs to change location and exclude the rival from the market.

A different setup must then be constructed to allow for a trade-off between locational choice and adoption choice. For instance, we might assume that the cost of adoption varies over space according to the firm's proximity to the location of the supply of the new technology [11]. Better information or the existence of specialized repairing services may cause an advantage in adoption for firms located closer to a market center where the innovation is offered [12].

In principle it would be possible to consider an alternative order for the adoption-location game described in the previous sections with firms choosing location (and price) at the first (and second) stage and adoption time at a later stage. In this case, firms should take into account, in choosing the initial location, successive decisions about adoption and locate closer or further according to the date at which they believe to adopt.

Hence it might be possible to reach an equilibrium in which one firm decides to adopt soon and locate close to the center of supply and the other delays adoption but captures a greater market share [13].

The inclusion of other elements in a spatial competition framework for the analysis of the strategic adoption of new technology in space is also possible. Although potential extensions will surely complicate the framework analytically they might be fruitful to solve some inexistence problems and to improve the connection with real world phenomena. The extension of the model to more than two firms and to leader-follower setups and the

---

[11] Incorporation of the supply side in spatial diffusion has been suggested by BROWN (1981) as the true "new perspective" of spatial diffusion modeling. Also, interaction between innovation demand and supply has been considered the most important improvement in current literature on innovation diffusion (STONEMAN, 1987).

[12] It must be stressed the difference between location-dependent production cost and location-dependent adoption cost. If we allow production cost to be different according to the location, some equilibrium problems can be avoided but tractability seems rather difficult (SCHULZ and STAHL, 1985).

[13] For instance, in a linear market, this case occurs when the center that supplies the innovation is located towards or at one extremity.

incorporation of R&D activity and of uncertainty in the adoption process seem to be quite promising avenues for research. In its current setup, the framework may only capture the rationale of adoptions occurring in the so-called "supplier dominated sector" (Pavitt, 1984) in which firms conduct limited R&D activity with minimal role for appropriability and cumulativity of technological change. Distribution networks in textiles could have the required scale economies in production to justify the duopoly (or oligopoly) setup. In addition, most of the content of the previous sections can find a better interpretation if referred to the product space rather than to a geographic location space. In the product differentiation case, however, price discrimination behavior can be more difficult to adopt. Nevertheless, the very recent advances in the theory of imperfect competition (Caplin and Nalebuff, 1991) may broaden considerably the scope of application of the setup presented. By offering existence and uniqueness results for the pure strategy price equilibrium for any number of (asymmetric) firms and for any location, we are open to a wide variety of applications of the subgame perfect equilibrium concept both in location and product differentiation contexts.

REFERENCES

BRAID R, "Heterogeneous Preferences and Non-Central Agglomeration of Firms", *Regional Science and Urban Economics*, 1988, *15*, 229-43.

BROWN L., *Innovation Diffusion. A New Perspective*, London and New York: Methuen, 1981.

BRESNAHAN T.F., "Duopoly Models with Consistent Conjectures", *American Economic Review*, 1981, *71*, 934-45.

CAPLIN A. and NALEBUFF B., "Aggregation and Imperfect Competition: On the Existence of Equilibrium", *Econometrica*, 1991, *59*, 25-59.

DE PALMA A., GINSBURGH V., PAPAGEORGIOU Y.Y. and THISSE J.F., "The Principle of Minimum Differentiation Holds under Sufficient Heterogeneity", *Econometrica*, 1985, *53*, 767-81.

EATON J. and KIERZKOWSKI H., "Oligopolistic Competition, Product Variety, Entry Deterrence and Technology Transfer", *Rand Journal of Economics*, Spring, 1984, 99-107.

FUDENBERG D. and TIROLE J., "Preemption and Rent Equalization in the Adoption of New Technology", *Review of Economic Studies*, 1985, *52*, 383-401.

GABSZEWICZ J.J. and THISSE J.F., "Spatial Competition and the Location of Firms", in J.J. Gabszewicz, J.F. Thisse, M. Fujita and U. Schweizer, *Location Theory*, London: Harwood, 1986.

HURTER A.P., and LEDERER P.J., "Spatial Duopoly with Discriminatory Pricing, *Regional Science and Urban Economics*, 1985, *15*, 541-53.

LEDERER P.J. and HURTER A.P., "Competition of Firms: Discriminatory Pricing and Location", *Econometrica*, May, 1986, *54*, 623-40.

PAVITT K., "Patterns of Technical Change: Towards a Taxonomy and a Theory", *Research Policy*, 1984, *13*, 343-73.

PRESCOTT E.C. and VISSCHER M., "Sequential Location among Firms with Foresight", *Bell Journal of Economics*, Autumn, 1977, 378-93.

REINGANUM J., "On the Diffusion of New Technology: A Game Theoretic Approach", *Review of Economic Studies*, 1981, *48*, 395-405.

SCHULZ N. and STAHL K., "On the Non Existence of Oligopolistic Equilibria in Differentiated Products Spaces", *Regional Science and Urban Economics*, 1985, *15*, 229-43.

STONEMAN P., *The Economic Analysis of Technological Policy*, Oxford: Clarendon Press, 1987

TIROLE J., *The Theory of Industrial Organization*, Cambridge: The MIT Press, 1988.

CONCORRENZA SPAZIALE E ADOZIONE DI NUOVA TECNOLOGIA

Il lavoro mostra come lo schema di concorrenza spaziale possa essere utilizzato per analizzare l'adozione strategica di nuova tecnologia in un caso di duopolio. Il caso di localizzazione prefissata viene posto a confronto con quello in cui le imprese possono modificare la loro localizzazione successivamente alla adozione di nuova tecnologia. Il saggio si sofferma sui problemi connessi all'esistenza dell'equilibrio nel caso di scelta sequenziale localizzazione-prezzo con costi asimmetrici. Alla fine del lavoro si suggeriscono alcune estensioni e limiti dell'approccio prescelto.

Rivista Internazionale di Scienze Economiche e Commerciali
Volume 39 (1992), N. 1, 49-67

ARE THERE REAL OR MONETARY BUSINESS CYCLES IN THE UNITED KINGDOM ECONOMY?

by
IOANNIS A. KASKARELIS *

I. *Introductions*

Business cycle theory was a major branch of economics until the keynesian revolution. Early business cycle theorists considered the cycle as largely self-sustained, where each boom contains the seeds of recovery and boom. The keynesian revolution shifted the focus of macroeconomics from the inevitability of the cycle to methods of improving macroeconomic performance. During the long expansion of the sixties, it was even possible to think that business cycle has been cured. In the decade which followed, poor economic performance (reflected in the increased frequency and depth of recessions) and the forceful advocacy of Lucas (1977) renewed interest in the business cycle as a specific field of research. Nevertheless, the eighties experienced a new challenge in business cycle theory. This affected all conventional views (keynesian and new-classical) since the definition of business cycle itself and the alleged size of the effects of macroeconomic policy (in particular monetary policy) on economic activity, is questioned. According to the new real business cycle (RBC) theory, the size and persistence of the cycles is much smaller than commonly thought. Therefore the cycle is much less of a problem for economic policy.

The distinguishing characteristic of RBC models is a denial that monetary policy actions have any significant impact on aggregate output and employment magnitudes (see McCallum, 1986; Fischer, 1988). Such models attribute all business cycle phenomena to changes in taste patterns and productivity (real) shocks. However, the RBC point of view does not deny

* Birkbeck College, University of London, London.
I would like to thank Nikitas Pittis, Prof. Ron Smith and an anonymous referee of this journal for useful comments and suggestions. Errors remain unerringly my own.

that there is any association between output and monetary magnitudes but it attributes the observed money-output correlation to the so-called 'reserve causation', i.e. responses of the money stock, via the monetary authority and/or the banking sector, to variations in aggregate output (see eg. King and Plosser, 1984). Thus, what the RBC theorists claim is that observed Phillips-type correlations stem from the monetary system's reaction to output fluctuations (the latter induced by real shocks to tastes or technology) and not from the nonbank private sector's reaction to monetary shocks.

The influence of this new point of view has been substantial because it offers a theory based on strict neoclassical principles at a time when the attractiveness of the leading alternative (the Barro/Lucas theory of monetary business cycles) seems to weaken. This induces a tremendous amount of research on a better specification of the role of monetary policy in a neoclassical model. The pioneer theoretical paper in the RBC theory was Kydland and Prescott (1982) which showed that several business cycle correlations can be mimicked reasonably well by a competitive equilibrium model in which neither money nor government policy plays any role whatsoever. Despite the fact that theoretical works have not managed to proceed significantly since then (see the discussion in Lucas, 1987), there is a growing number of empirical papers which try to identify the significance of monetary shocks on real variables. First, there is a line of argument developed primarily by Nelson and Plosser (1982) that relies entirely on the univariate time-series properties of aggregate output, employment and other real variables. The main emphasis of their argument is that most of the fluctuations in these variables should be attributed to the trend component, in a secular versus cyclical decomposition, which would presumably be unaffected by monetary shocks. On the other hand, there are the studies of Sims (1980, 1982), Litterman and Weiss (1985) which show that money 'stock innovations' explanatory power for output variations diminishes when some interest rate variables are included in the Vector Autoregression (VAR) system.

There is a relative aversion among British economists in using these 'American style', atheoretical econometric methods in order to investigate the real sector of the economy. Empirical evidences for the UK economy are presented in cross-country studies, which are usually works of continental or American researchers [1]. British researchers are rather devoted to an

[1] See e.g. STULZ and WASSERFALLEN (1985) for unit root examinations on four (*GNP*, industrial production, unemployment rate, and real wages) UK series, AHMED et al. (1989) for tests on several VAR models; even BLANCHARD and SUMMERS (1986) for testing the existence

effort of finding a reasonable explanation for the collapse (since the early 70s) of demand for money relationship (see Goodhart, 1989). Trying to achieve that, they use econometric methods and well defined statistical models which incorporate the long established error correction mechanism.

The purpose of this paper is to discuss the relevance and some of the propositions of the real business cycle theory in the context of the UK economy, for the period 1963: I to 1987: IV. During this period, real disturbances of great magnitude (like the oil price increases, large changes in the real exchange rates and in fiscal policy) occurred. These phenomena had an impact on output and employment. Section II investigates the existence of unit roots for sixteen quarterly macroeconomic series in the UK. In the econometric specification of the hypotheses tested, I tried to avoid using any tricky method of detrending and separating the secular from the cyclical component. Section III examines several VAR models and tests for causality among output and several real and nominal variables. Here I take into consideration the problem of lag structure specification. Additionally I check for cointegrated series among them. Section IV sums up the conclusions.

II. *Trend and Cyclical Components – Unit Roots in Macroeconomic Series*

It is a common practice in theoretical and empirical works on business cycles, to correct the time-series under investigation for the growth component, in order to isolate the cyclical part. The unobserved components model is of the form

$$Y_t = \bar{y}_t + c_t$$

where c_t is the 'cyclical' component supposed generated by a process that has the property of stationarity, and $\bar{y}_t$ is the 'secular' component to which any non-stationarity in y_t must be attributed.

Two approaches are used to account for secular movements in empirical studies of the business cycle:

(*i*) In the first approach, the observable variables which are considered for economic growth are explicitly included in the respective regressions. E.g. the model could contain a deterministic linear time trend *plus* a stationary stochastic component with an unconditional mean of zero however of

of unit roots in unemployment series, or GRILLI (1988) for using integration and cointegration tests investigating the money causality and the debt-segniorage relation.

unspecified autocorrelation structure. The former variable approximates the secular component while the latter is taken as an adequate measure of c_t. This class of models is the so-called Trend Stationary (TS) models. It is important to note here that the secular part is of a deterministic nature and only the cyclical component is stochastic. Proponents of RBC theory state that the assumption of a deterministic trend is ad hoc. Nelson and Plosser (1982) argue that using observable variables to account for growth components seems unsatisfactory since neither factor inputs nor population seem to suffice and direct measures of technology are not readily available.

(*ii*) Nelson and Plosser (1982) assume that y_t is generated by a 'difference stationary' (DS) model. Thus, y_t is a variable whose ARMA representation includes a unit root in the AR polynomial and no deterministic trend. Assuming that c_t is stationary it then follows that the secular component must have a unit root. Analysing the variance of the components, they conclude that the variance in actual output changes for the USA is dominated by changes in the secular component $\bar{y}_t$ rather than the cyclical one c_t. The weak point in their study is the crucial role of conclusions regarding DS versus TS processes. The cyclical component is measured by whatever is left over after an estimate of the (DS) secular component (practically random walk) is removed. However, if the process is of a TS type with an AR root close to one, then the secular component removal step can easily take out many times as much of the signal as is properly attributable to the secular component, thereby yielding a many-fold underestimate of cyclical variability.

A proper way to avoid the detrending problems that Nelson and Plosser (1982) faced is to consider a model that encompasses the TS and DS process as special cases:

$$\Delta y_t = \alpha + \beta t + \rho y_{t-1} + u_t \tag{1}$$

where $\Delta y_t = y_t - y_{t-1}$, t the time trend and u_t the error term.

(a) $\beta \neq 0$ and $\rho = -1$ implies a TS model discussed previously.

(b) $\beta = 0$ and $-1 < \rho < 0$ is an ARMA Box/Jenkins class of models.

(c) $\beta = 0$ and $\rho = 0$ is a DS model where y_t variable is integrated of order one: I (1). Assuming that the cyclical component is stationary, the secular component has a unit root and y_t follows a random walk process. In this case future values of y_t cannot be predicted since the disturbance term is not autocorrelated. Furthermore if $\alpha \neq 0$ (random walk with drift process) the forecast variance increases 'without bound' as the series is an accumulation of stationary changes, and this sum is not stationary.

Dickey and Fuller (1979) provide tabulations of the distributions of the t-ratio for ρ [Appendix A analyzes the statistical problems arising when a series is integrated and presents the statistical method that Dickey and Fuller have introduced]. In a later paper Dickey and Fuller (1981) construct likelihood ratio tests for the joint null hypothesis ϕ_2: $(\alpha, \beta, \rho) = (0, 0, 0)$ and the ϕ_3: $(\alpha, \beta, \rho) = (\alpha, 0, 0)$ [2]. Their estimated values are compared with the values of ϕ_i^* ($i = 2, 3$) provided in the same paper: if $\phi_i < \phi_i^*$ the null hypothesis is accepted.

One problem arises when u_t is autocorrelated. One way to handle it is to introduce parametric approximations, like the Augmented Dickey Fuller (ADF) test, to the process generating the disturbance term. The ADF test is carried out with the OLS estimation of

$$\Delta y_t = \alpha + \beta t + \rho y_{t-1} + \sum_{1}^{k} y_i \Delta y_{t-i} + u_t \tag{2}$$

and checks if ρ is significantly different from zero. Comparing the two formulas (1) and (2) one sees that the role of the $\Sigma y_i \Delta y_{t-i}$ term is to 'soak up' the serial correlation of error term. If for instance Δy_t follows a stationary AR (p) process with p known then the null hypothesis of a unit root can be tested, by estimating an autoregression of Δy_t on its p lags and y_{t-1}. Said and Dickey (1984) extended the method to the case that Δy_t follows a generalized ARMA (p, q) process with p, q unknown: a regression model such as (2) is still valid, if the number of lags k of Δy_t introduced as regressors increases with the sample size at a controlled rate of $N^{1/3}$, where N is the number of observations. It must be noticed that only if u_t is a white noise can the limiting distributions of the test statistics be obtained (see Perron, 1988). Thus in (2), the first differenced series are white noises only if all the estimated coefficients of lagged Δy_{t-i} are statistically zero.

There are sixteen macroeconomic series tested for unit roots during the period 1963: I – 1987: IV [3]. All variables except the unemployment rate

[2] In order to perform the tests the following regressions have been used

$$\Delta y_t = \nu \Delta y_{t-1} + u_t \tag{2a}$$

$$\Delta y_t = \alpha + \nu \Delta y_{t-1} + u_t \tag{2b}$$

Thus, for the ADF model (SSR(i) is the estimated sum of squared residuals of the (i) equation)

$$\phi_2 = \frac{SSR\,(2a) - SSR\,(2)}{SSR\,(2)} \frac{N-8}{7}, \text{ and } \phi_3 = \frac{SSR\,(2b) - SSR\,(2)}{SSR\,(2)} \frac{N-8}{6}$$

[3] National account variables are from the OECD "Quarterly National Accounts", various issues. The rest of them are from the OECD "Main Economic Indicators". All the series are

and the Treasury Bill rate are in logarithms. Experiments with TS models showed that the cyclical component, measured as the deviation from a deterministic trend is highly and positively autocorrelated. The DS model and Box-Jenkins analysis showed that business cycles do not exceed the 5th quarter and furthermore the cyclical part is rather approximated by low order MA processes, which coincide with the Nelson and Plosser (1982) and Stulz and Wasserfallen (1985) type of results. We should also point out here that estimates on the simple DF model (1) revealed that the hypothesis of white noise residuals is highly rejected in the twelve out of sixteen cases, which implies that we can not state it as the appropriate model [4].

Table 1 presents the estimated coefficients for $\hat{\alpha}$, $\hat{\beta}$, $\hat{\rho}$ and the relevant ADF of equation (2). There are also reported the standard error of the regression (s), Durbin-Watson statistic (DW), Box-Pierce Q-statistic for 20 lags (asymptotically distributed by χ^2 (20)) and two parameter stability tests, with break points: (a) the 1973: IV quarter (first oil shock) and (b) the 1981: I quarter (switch to a monetarist economic regime by the conservative government). In case that the hypothesis is rejected, tests for unit root are applied for the different sub-samples (Table 2).

Results for the ADF test show that the hypothesis for white noise residuals is accepted in all cases at 5% s.l.. All series seem to follow random walk with drift processes. The ϕ_3 hypothesis is accepted at 5% in all cases except that of total employment where it is accepted at 10%. ϕ_2 hypothesis is accepted at 5% in eight cases however clearly rejected in four. The trend parameter β becomes insignificant for the majority of the regressions as judged by the usual t-statistic. Q-statistic, standard error and Durbin-Watson indicate no further problems. Regressions for different sub-samples (Table 2) indicate that the main results coming out from Table 1 do not change significantly. The only exception is represented by the employment series in the earlier period whose results show sluggishness in adjustment to long run equilibrium, rather than existence of unit root. The previous findings support the Stulz and Wasserfallen (1985) line of argument, that a modern theory of macroeconomic fluctuations must explain changes in the trend of output and deviations from the trend of output.

Conclusions for these UK macroeconomic series are fairly similar to those of US annual series examined by Nelson and Plosser (1982). They also coincide with those by Stulz and Wasserfallen (1985) for the correspond-

---

seasonally adjusted except the Treasury Bill Rate. This is plausibly the reason why the seasonal component has been omitted from the above discussion on the variable's decomposition.

[4] The above mentioned results are available upon request by the author.

TABLE 1

ADF TEST FOR UNIT ROOTS: EQUATION (2)

	$\hat{\alpha}$	β	$\hat{\rho}$	ϕ_2	ϕ_3	Q-stat.	s	D-W	Chow Tests 73:IV	Chow Tests 81:I
Nominal GDP										
	0.051	0.001	– 0.028	10.42	2.64	11.39	0.02	2.02	3.55	1.26
	(2.21)	(1.73)	(– 1.75)							
Real GDP										
	0.493	0.001	– 0.133	4.60	1.53	14.09	0.01	2.00	1.30	0.86
	(2.62)	(2.30)	(– 2.55)							
Nominal Domestic Demand										
	0.050	0.001	– 0.028	13.14	4.10	11.33	0.02	2.04	2.48	1.00
	(2.32)	(1.87)	(– 1.88)							
Real Domestic Demand										
	0.505	0.001	– 0.137	3.46	1.61	16.97	0.01	2.03	1.96	0.58
	(2.88)	(2.68)	(– 2.84)							
Real Private Consumption										
	0.432	0.001	– 0.138	6.13	2.86	16.13	0.01	2.00	0.95	1.15
	(2.42)	(2.45)	(– 2.39)							
Real Fixed Investment										
	0.231	0.000	– 0.106	2.92	1.70	7.19	0.03	1.98	0.72	1.28
	(2.89)	(1.64)	(– 2.72)							
Real Exports										
	0.160	0.001	– 0.074	3.28	1.28	11.06	0.04	1.98	1.65	0.37
	(1.58)	(1.08)	(– 1.28)							
Real Imports										
	0.286	0.001	– 0.144	2.19	1.28	8.35	0.04	2.02	2.10	1.35
	(2.32)	(2.13)	(– 2.19)							
Total Employment										
	1.282	0.000	– 0.128	5.63	6.57	17.26	0.03	2.03	2.44	3.28
	(1.90)	(1.02)	(– 1.90)							
Unemployment Rate										
	– 0.001	0.000	– 0.040	1.70	1.99	9.20	0.00	1.97	0.69	2.62
	(– 0.78)	(2.12)	(– 2.34)							
GDP Deflator										
	– 0.038	0.000	– 0.019	6.57	2.79	11.28	0.01	1.95	2.15	1.29
	(– 1.94)	(2.19)	(– 2.17)							
Domestic Demand Deflator										
	– 0.035	0.000	– 0.017	5.63	4.38	9.41	0.01	1.96	1.75	1.08
	(– 2.16)	(2.34)	(– 2.38)							
Treasury Bill Rate										
	0.011	0.000	– 0.180	1.52	1.57	11.79	0.01	1.95	1.00	2.50
	(2.32)	(1.15)	(– 2.77)							
M1 Money Supply										
	0.266	0.001	– 0.033	9.00	3.83	18.56	0.02	2.00	0.56	1.53
	(1.73)	(2.21)	(– 1.73)							
Real Money (M1/Domestic Demand Deflator)										
	0.177	0.000	– 0.179	3.00	3.23	13.95	0.02	2.00	1.72	5.10
	(0.58)	(1.66)	(– 0.62)							
Velocity (M1/Nominal Domestic Demand)										
	– 0.053	0.000	0.005	2.00	2.00	13.53	0.03	1.99	1.41	3.60
	(– 0.23)	(1.27)	(0.16)							

Critical Value:

sl/st	$\hat{\tau}_{\rho}$	ϕ_2	ϕ_3
10%	2.38	4.16	5.45
5%	2.79	4.88	6.49
1%	3.53	6.50	8.73

Note: *t*-statistic in parenthesis. Chow-test is an *F* (8,84) or *F* (8,76) test.

TABLE 2

ADF TEST FOR UNIT ROOTS UNDER DIFFERENT SAMPLES

sample	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$	ϕ_2	ϕ_3	Q-stat.	s	D-W
Nominal GDP								
63:I-73:IV	– 0.179	– 0.002	0.119	7.2	1.2	7.91	0.02	2.00
	(– 0.93)	(– 0.72)	(1.02)					
74:I-87:IV	0.138	0.000	– 0.034	9.3	5.1	9.28	0.01	2.12
	(3.78)	(0.40)	(– 1.37)					
Nominal Domestic Demand								
63:I-73:IV	– 0.137	– 0.001	0.086	6.5	1.1	7.07	0.02	1.96
	(– 0.64)	(– 0.47)	(0.67)					
74:I-87:IV	0.139	0.001	– 0.045	9.8	5.7	12.04	0.01	2.03
	(3.52)	(1.09)	(– 1.91)					
Real Domestic Demand								
63:I-73:IV	1.562	0.003	– 0.435	3.2	1.3	10.89	0.02	2.02
	(2.64)	(2.54)	(– 2.61)					
74:I-87-IV	0.344	0.001	– 0.096	2.5	2.0	9.99	0.01	2.09
	(1.32)	(2.02)	(– 1.38)					
Real Imports								
63:I-73:IV	0.581	0.005	– 0.327	3.0	0.7	7.00	0.03	1.99
	(1.21)	(1.30)	(– 1.18)					
74:I-87:IV	0.288	0.002	– 0.158	3.4	2.7	7.75	0.04	2.03
	(1.23)	(2.03)	(– 1.41)					
Total Employment								
63:I-73:IV	5.299	– 0.001	– 0.525	5.9	6.8	10.32	0.02	2.02
	(2.68)	(– 2.13)	(– 2.68)					
74:I-87:IV	1.748	0.001	– 0.178	1.2	1.1	6.25	0.03	1.90
	(1.94)	(1.87)	(– 1.96)					
63:I-81:I	3.311	– 0.000	– 0.329	7.3	8.6	24.26	0.02	2.04
	(2.49)	(– 1.59)	(– 2.49)					
81:II-87:IV	3.736	0.003	– 0.404	1.1	1.3	2.25	0.03	2.04
	(2.33)	(– 2.14)	(– 2.35)					
Unemployment Rate								
63:I-81:I	– 0.000	0.000	– 0.044	3.0	3.3	8.48	0.00	1.97
	(– 0.91)	(1.92)	(– 1.26)					
81:II-87:IV	0.045	– 0.000	– 0.144	1.2	1.4	5.96	0.00	1.95
	(2.25)	(– 1.89)	(– 1.64)					
GDP Deflator								
63:I-73:IV	– 0.094	0.001	– 0.051	4.9	3.1	8.98	0.01	1.55
	(– 1.07)	(1.68)	(– 1.07)					
74:I-87:IV	0.023	– 0.000	– 0.011	5.2	1.6	11.18	0.01	1.70
	(0.48)	(– 0.20)	(– 0.66)					
Treasury Bill Rate								
63:I-81:I	0.011	0.001	– 0.432	3.1	3.5	8.74	0.01	2.02
	(2.51)	(3.34)	(– 3.82)					
81:II-87:IV	0.046	– 0.000	– 0.265	1.4	1.6	11.16	0.01	1.78
	(0.69)	(– 0.43)	(– 1.21)					
Real Money (M1/Domestic Demand Deflator)								
65:I-81:I	2.893	– 0.001	– 0.274	2.6	3.0	10.67	0.02	2.01
	(3.84)	(– 2.80)	(– 3.84)					
81:II-87:IV	– 0.379	0.002	0.021	6.3	4.7	7.01	0.02	2.05
	(– 0.46)	(0.80)	(0.21)					
Velocity (M1/Nominal Domestic Demand)								
65:I-81:I	2.391	– 0.002	– 0.347	2.6	2.0	13.34	0.02	2.04
	(3.44)	(– 3.36)	(– 3.47)					
81:II-87:IV	– 0.510	0.002	0.049	2.7	3.2	7.58	0.02	1.94
	(– 0.65)	(1.16)	(0.33)					

ing GB monthly series for the same period. The rejection of the null hypothesis for the pre-seventies unemployment rate in their study, is reflected in a near failure in the unemployment rate and total employment series for the 1963-73 period in this study. Time trend is also significant in the pre-seventies and in the eighties sub-samples. The constant term is significantly negative for the whole sample but significant and positive for the 1980's subsample, which is not examined in the Stulz and Wasserfallen (1985) study.

An interesting issue coming out from the results is that unemployment rate series appears to follow a random walk. This result differs from that of Nelson and Plosser's (1982) for the United States, where the unemployment rate is the only series which does not seem to follow a random walk. In the case of the United Kingdom this characteristic suggests that unemployment does not follow a pronounced cyclical pattern but instead shows shifts which are persistent. This is particularly true for the seventies and early eighties when unemployment increased to a new level and stayed there. In economic literature we have seen a number of possible explanations for this phenomenon (e.g. the hysteresis hypothesis, see Blanchard and Summers, 1986).

III. *Vector Autoregression Systems and Granger-Causality Tests*

The test procedures used above have virtually no power in discriminating between a first order unit root and an autocorrelation coefficient slightly below zero. However the economic interpretation is completely different in the two cases. If a unit root is present the series is non-stationary, whereas in the other case it would return to a constant long-run mean (see e.g. McCallum, 1986). Thus, it is worth analyzing further the aspect of real business cycle theory by looking at the effects of monetary and real variables on the development of output. For this purpose I am going to use Vector Autoregression models. VAR proved to be a convenient way of summarizing empirical regularities and perhaps suggesting predominant channels through which relations work (see Sims, 1982).

Of the various VAR studies, the first to appear was that of Sims (1980) which was followed by Sims (1982). In these papers Sims estimates VAR systems that include among their variables measures of aggregate production and money stock. He solves for the implied moving-average representations and uses the latter to decompose the variance of each variable into portions attributable to the innovations of each of the system's variables. He finds for the USA that when a system includes only money output and the price level, the money stock innovations contribute a substantial fraction of the total explanatory power for output. However when some nominal interest rate is added to such system, the fraction of output variability attributable to money stock innovations declines sharply. He concludes that monetary policy surprises are not important in explaining the real component of postwar business cycles. Thus, the imposition of a monetarist rule to make the quantity of money more predictable would have had little real effect towards reducing these fluctuations [5]. Ahmed et al. (1989) confirm Sim's conclusion for the European Community. Litterman and Weiss (1985) also find that the portion of output variance attributable to money stock innovations declines sharply when a nominal interest rate is added to a small VAR system. However one of their prominent finding is that in the quarterly US data the real rate is not significantly Granger-caused by any of money, output, price and the nominal interest rate. They argue that the theories of Lucas/Barro and sticky-price types are contradicted by the data, as both transmit monetary impulses to real variables by way of the real rate. Other results reported in Litterman and Weiss (1985) provide evidences which are inconsistent with the RBC hypothesis. Figures in their paper indicate that log of output is in fact Granger-caused by nominal variables.

In the paragraphs that follow I am going to test some of the above propositions for the United Kingdom economy. The variable representing the measure of output is domestic demand (real *GDP* minus net exports), denoted as *D*. As far as monetary factors are concerned, we examine the effectiveness of *M*1 money supply, *M*, and the role of short-term interest rate (Treasury Bill rate), *I*. In further stages we will also examine the real money supply (*M*1 minus domestic demand deflator), *RM*, and the real interest rate (Treasury Bill rate minus expected quarterly rate of domestic expenditure price inflation), *RI*. The real variables, which reflect influences from the real side of the economy, are the terms of trade (imports over exports deflator) *TR*, exports *X*, government expenditures *G*, and real production wages *W*. Therefore we could investigate the importance of the oil price shocks, international macroeconomic interdependence, government activity, and the consequences of trade unions' actions. All variables except the interest rates are first differences of natural logs. Interest rates are simple differences. VAR models assume that stochastic processes are stationary.

[5] Nevertheless money stock innovations do not represent the sole component of monetary policy actions (see e.g. KING and PLOSSER, 1984).

Consequently, first differences Δy_t have been used in the light of evidences coming out from the previous section.

VAR estimation pays particular attention to determining the appropriate lag structure. In Sims (1980, 1982) and Litterman and Weiss (1985) all variables appear in all equations with the same number of lags. In this paper I am going to follow the procedure proposed by several authors (see Hsiao, 1981; Fuckler, 1985; Ahmed et al., 1989; Scheide, 1989, among others), which implies that the lag length in the vector autoregressions should be chosen according to Akaike's FPE creterion (Final Prediction Error). This is more preferable than the usual practice of ad-hoc lags. In that method results could be biased because either existing causalities are not detected or spurious causalities. [Appendix B analyzes the FPE procedure]. Table 3 reports the optimum lag length and the corresponding FPE for each univariate case. Before proceeding to the construction of the bi- or multivariate VAR models we must be sure that important error correction terms are not erroneously excluded from the regressions. This examines the importance of long run components in regressions among difference stationary variables, and it is connected with the rapidly expanding literature on cointegration. [Appendix C examines the importance of cointegration and the tests for cointegration]. The results in Table 4 show that the null non-cointegration hypothesis cannot be rejected at 5% s.l. in all cases except those of the bivariate systems between output and real wages. Hence, in the causality tests that follow an error correction term will be included in the [ΔD; ΔW] and [ΔW; ΔD] VAR systems, according to the Engle and Granger (1987) two-step procedure. Table 5 investigates what Granger-causes ΔD, whereas in Table 6 the inverse causality question is applied. In all equations the tests for autocorrelation in the residuals (calculated from the *Q*-statistics) are favourable to the hypothesis of white noise processes. Estimates show that real domestic demand is caused by money supply and competitiveness. On the other hand demand causes money, nominal interest rate (in the limit), competitiveness and real wages [6]. In the [ΔD; ΔM], [ΔD; ΔRM] and [ΔD; ΔTR] systems causality runs in both directions; neither the supply nor the demand determined output approach could be rejected. Real wages exhibit procyclical behavior and they are affected by demand changes. Estimates on the [ΔD; ΔW] system indicate that a long run non-accelerating inflation natural rate of output exists for the UK economy. Another interesting point coming out is that

[6] Demand also causes the real money supply, and is caused by the real money supply and the real interest rate.

TABLE 3

OPTIMAL LAG LENGTH FOR THE UNIVARIATE AUTOREGRESSIONS

	opt. lag			Chow-Tests: Split at:	
series	length	FPE * 10^3	Q-stat.	73:IV	81:I
ΔD	1	0.213	16.77	1.7 (2,92)	1.0 (2,92)
ΔM	4	0.486	17.81	1.1 (5,85)	2.6 (5,85)
ΔI	1	0.191	17.45	0.3 (2,92)	0.5 (2,92)
ΔTR	3	0.094	15.34	0.3 (4,88)	0.5 (4,88)
ΔG	1	0.187	16.01	0.5 (2,92)	0.8 (2,92)
ΔX	1	1.390	17.85	1.1 (2,92)	0.1 (2,92)
ΔW	3	0.234	11.21	1.1 (4,88)	0.9 (4,88)

Note: The ΔM VAR for the sample 81:II-87:IV give opt. lag length 4; FPR * $10^3 = 0.434$, Q-stat. $= 7.30$.

TABLE 4

TESTS FOR COINTEGRATION

VAR system	$\hat{\delta}$-coefficients	CRDW	DF	ADF
[D, M]	0.16	0.13	2.50	2.30
[D, I]	2.74	0.13	0.89	0.56
[D; TR]	1.69	0.30	0.67	0.53
[D; G]	0.97	0.20	1.44	2.16
[D; X]	0.41	0.44	3.01	1.93
[D; W]	0.93	0.56	4.07	3.08
[M; D]	5.51	0.12	2.16	1.93
[I; D]	0.14	0.32	2.60	2.12
[TR; D]	0.08	0.13	1.80	1.91
[G; D]	0.95	0.20	1.67	2.27
[X; D]	2.30	0.45	3.15	2.06
[W; D]	1.05	0.56	4.03	3.03
[M; I]	15.01	0.11	0.61	0.15
[I; M]	0.02	0.32	2.69	2.23
[D; M, I]	0.15, 0.41	0.14	2.33	2.14
[M; D, I]	5.55, – 0.23	0.12	2.16	1.93
[I; M, D]	– 0.002, 0.15	0.32	2.59	2.12
[M; TR]	6.84	0.01	0.69	0.22
[TR; M]	0.01	0.13	1.74	1.90
[D; M, TR]	0.16, 0.61	0.17	2.83	2.48
[M; D, TR]	5.74, – 2.89	0.15	2.48	2.09
[TR; D, M]	4.48, – 3.42	0.17	2.15	2.11
			critical value	
bivariate system (5% s.l.)		0.386	3.37	3.17
Multivariate system (5% s.l.)			4.22	4.02

Note: In the VAR system presentation the semi colon, e.g. [X; Y; Z], separates the dependent variable X from the independent ones Y and Z.

Table 5

WHAT CAUSES REAL DOMESTIC DEMAND?

opt. lag					Chow Tests: Split at:	
series	length	FPE * 10^3	cause?	Q-stat.	73:IV	81:I
ΔM	1	0.208	yes	15.39	2.6 (3,89)	0.1 (3,89)
ΔI	1	0.217	no	15.98	1.1 (3,90)	0.3 (3,90)
ΔTR	4	0.212	yes	11.89	1.2 (6,84)	0.8 (6,84)
ΔG	1	0.217	no	16.64	1.8 (3,90)	0.9 (3,90)
ΔX	1	0.217	no	16.90	1.7 (3,90)	0.7 (3,90)
ΔW	1	0.221	no	16.77	1.1 (4,89)	0.8 (4,89)

Table 6

WHAT IS CAUSED BY REAL DOMESTIC DEMAND?

DD opt. lag					Chow Tests: Split at:	
series	length	FPE * 10^3	cause?	Q-stat.	73:IV	81:I
ΔM	3	0.449	yes	15.66	0.4 (8,79)	1.2 (8,79)
ΔI	2	0.191	yes	11.79	0.8 (4,88)	2.5 (4,88)
ΔTR	4	0.093	yes	12.94	0.2 (8,80)	1.2 (8,80)
ΔG	1	0.188	no	18.41	0.1 (3,90)	0.4 (3,90)
ΔX	1	1.410	no	19.54	2.0 (3,90)	0.2 (3,90)
ΔW	2	0.226	yes	15.11	0.4 (7,83)	0.6 (7,83)

domestic demand neither causes nor is caused by government expenditure and exports, while Chow-tests did not confirm any parameter instability problem between different sub-samples.

Table 7 shows the direction of causality among ΔD, ΔTR, ΔM, ΔI, ΔRM and ΔRI that is coming out from several bi- and multivariate systems [7]. Granger causality between interest rate and money runs to both directions. However the effect of money on output dominates over that of nominal interest rate. This means that money continues to play an important role in explaining output movements even when an interest rate is included in the system. Although the revealed direction of causality between ΔM and ΔI is surprising, it can be argued that actions of the central bank, via changes in interest rates or in money supply, obviously affect output. The next set of relationships reveals lack of causality in any direction between money and competitiveness, but the former seems to dominate over the latter in caus-

[7] All regressions analyses are available upon request by the author.

TABLE 7

CAUSALITY DIRECTION FROM SEVERAL VAR SYSTEMS

$\Delta I \longrightarrow \Delta M$ $\Delta I \not\rightarrow [\Delta D;\ \Delta M]$ $\Delta D \not\rightarrow [\Delta I;\ \Delta M]$	$\Delta M \rightarrow \Delta I$ $\Delta M \longrightarrow [\Delta D;\ \Delta I]$ $\Delta D \longrightarrow [\Delta M;\ \Delta I]$
$\Delta TR \not\rightarrow \Delta M$ $\Delta TR \not\rightarrow [\Delta D;\ \Delta M]$ $\Delta D \longrightarrow [\Delta TR;\ \Delta M]$	$\Delta M \not\rightarrow \Delta TR$ $\Delta M \longrightarrow [\Delta D;\ \Delta TR]$ $\Delta D \longrightarrow [\Delta M;\ \Delta TR]$
$\Delta I \not\rightarrow \Delta TR$ $\Delta I \not\rightarrow [\Delta D;\ \Delta TR]$ $\Delta D \not\rightarrow [\Delta I;\ \Delta TR]$	$\Delta TR \not\rightarrow \Delta I$ $\Delta TR \not\rightarrow [\Delta D;\ \Delta I]$ $\Delta D \longrightarrow [\Delta TR;\ \Delta I]$
$\Delta RM \not\rightarrow \Delta RI$ $\Delta RM \longrightarrow [\Delta D;\ \Delta RI]$ $\Delta D \longrightarrow [\Delta RM;\ \Delta RI]$	$\Delta RI \longrightarrow \Delta RM$ $\Delta RI \not\rightarrow [\Delta D;\ \Delta RM]$ $\Delta D \not\rightarrow [\Delta RI;\ \Delta RM]$
$\Delta TR \not\rightarrow \Delta RM$ $\Delta TR \not\rightarrow [\Delta D;\ \Delta RM]$ $\Delta D \not\rightarrow [\Delta TR;\ \Delta RM]$	$\Delta RM \longrightarrow \Delta TR$ $\Delta RM \longrightarrow [\Delta D;\ \Delta TR]$ $\Delta D \longrightarrow [\Delta RM;\ \Delta TR]$
$\Delta TR \not\rightarrow \Delta RI$ $\Delta TR \not\rightarrow [\Delta D;\ \Delta RI]$ $\Delta D \longrightarrow [\Delta TR;\ \Delta RI]$	$\Delta RI \longrightarrow \Delta TR$ $\Delta RI \not\rightarrow [\Delta D;\ \Delta TR]$ $\Delta D \not\rightarrow [\Delta RI;\ \Delta TR]$

Note: $\Delta y_1 \longrightarrow \Delta y_2$ means Δy_1 causes Δy_2. $\Delta y_1 \not\rightarrow \Delta y_2$ means Δy_1 does not cause Δy_2.

ing domestic demand. We have also lack of causality in the third set between competitiveness and interest rate, although ΔTR dominates over ΔD in affecting ΔI. This finding could be attributed to direct effects from exchange rates to interest rates.

The next sets of relationships include only real variables. The examination of real money and real interest rate come along with the Scheide (1989) thesis that only real variables should be used in VAR systems. He argues that "... although it is true that VAR can be viewed as reduced forms of a variety of structural models they are not immune to the Lucas-critique. But by using real money and interest rate we come relatively close to shocks of monetary policy". Real interest rate causes real money but the latter dominates over the former in affecting demand. Conclusively, results (either with real or nominal interest rate and money) coincide with those of the studies mentioned above in that money affects output, but are different in that interest rates dominate over money in causing output.

The last set of relationships is interesting as well. Real interest rate causes relative import prices but it does not dominate over them in causing domestic demand. On the other hand demand dominates over real interest rate (as over nominal interest rate) in causing competitiveness, but competitiveness dominates over demand in causing the real (as the nominal) interest rate probably through the exchange rate effects. Finally real money dominates over relative prices in all sets of equations.

Summing up, the results coming from the VAR models we have examined are in favour of the view of monetary business cycles. Money supply shows a strong causal link with output, while interest rates and terms of trade are also important but their relevance is sharply reduced or even vanishes if the money supply is taken into account. Nevertheless, we should point out here that it is hard to distinguish real effects from those stemming out from monetary policy since it is casually observed that positive supply shocks coincided with (or led to) monetary expansion. Equivalently, real shocks with negative consequences to output and employment were accompanied by monetary restrictions.

IV. *Synthesis*

This paper aimed at empirically investigating the argument of the real business cycle theory for the UK economy. For this purpose random walk tests and vector autoregressions have been used, in order to account for the evidence usually put forward by those who stress the dominance of real factors for fluctuations of economic activity. Results are not conclusive. The hypothesis that macroeconomic time series follow a random walk have found some support but VAR models showed that money plays a major role in output movements. Thus, since the view that monetary policy is ineffective could not be validated, the idea of monetary prescriptions cannot be dismissed. In particular a rule for monetary policy would help to stabilize the development of output (see Goodhart, 1989).

APPENDIX

The Issue of Integration

There is a growing awareness of the fact that economic data are rarely stationary, and that the particular type of non-stationarity observed has profound effects upon statistical inference. In many instances estimators that are traditionally taken to be normally distributed no longer are so if variables that are 'integrated' appear in the model (see Pagan and Wickens, 1989). According to Engle and Granger (1987) a series with no deterministic component which has a stationary, invertible Autoregressive Moving Average (ARMA) representation after differencing d times, is said to be integrated of order d and is denoted $I(d)$. Most economic time series are either $I(0)$ or $I(1)$, i.e. stationary in levels or in first differences. To illustrate the problem let us consider an AR (1) model

$$\Delta y_t = \rho y_{t-1} + u_t$$

where y_t and u_t are integrated of order zero ($\rho \neq 0$ and u_t a white noise process) random variables. Then it is generally the case that the sample means converge to their expectations and the $T(\hat{\rho})$ Student distribution of $\hat{\rho}$ is asymptotically normal. However, when the process is $I(1)$ the sample means converge to random variables and $T(\hat{\rho})$ is asymptotically the ratio of two random variables. The first practical solution to the problem was provided by Fuller (1976), Dickey and Fuller (1979, 1981). They show that in this case the statistical analogous to the regression t-statistic for the test of the hypothesis that $\rho = 0$ is

$$\tilde{\tau} = \hat{\rho}\,(S_u^2 c)^{-1/2}$$

Its asymptotic distribution is not normal but is the 'non-standard' Dickey-Fuller distribution tabulated by Fuller (1976). Compared with the normal distribution it is skewed strongly to the left making the normal a very poor approximation. S_u^2 is the regression residual mean square:

$$S_u^2 = (N-2)^{-1}\,[Y_t'\,(I - Y_{t-1}\,(Y_{t-1}'\,Y_{t-1})^{-1}\,Y_{t-1}')\,Y_t]$$

where N is the number of observations, Y are $(N-1)$ dimensional vectors and c is the lower-right element of $(Y_{t-1}'\,Y_{t-1})^{-1}$.

Dickey and Fuller provide a set of results that allow us to test the DS hypothesis against the TS hypothesis, as long as we are willing to assume that only AR terms are required to obtain satisfactory representations. The TS and DS processes are both special cases of

$$\Delta y_t = \alpha + \beta t + \rho y_{t-1} + u_t$$

for which we can test the hypothesis $\rho = 0$, $\beta = 0$. To prove that, let us consider the model

$$Y_t = a + bt + \frac{u_t}{(1-(1-\rho)L}$$

where L is the lag operator. Equivalently, after multiplying by $1-(1-\rho)L$

$$\Delta y_t = \rho y_{t-1} + [a\rho + (1-\rho)\,b] + b\rho t + u_t$$

If the TS hypothesis is correct then $\rho < 0$. If the DS hypothesis is correct then $\rho = 0$ and $\Delta y_t = b + u_t$.

Dickey and Fuller (1979) provide tabulations of the distributions of the t-ratio for ρ (denoted $\hat{\tau}_\rho$), for testing the null hypothesis $\rho = 0$:

$$\hat{\tau}_\rho = \hat{\rho}\,(S_u^2 c)^{-1/2}$$

In this case the regression residual mean square is given by

$$S_u^2 = (N-4)^{-1}\,[Y_t'\,(I - U\,(U'\,U)^{-1}\,U')\,Y_t]$$

where matrix U consists of the $(N-1)$ dimensional vectors $(1, t, Y_{t-1})$, c is the lower-right element of $(U'U)^{-1}$ and N the number of observations. They state that the distributions of $\hat{\rho}$ and $\hat{t}_\rho$ are not affected by whether α is zero or not but $\hat{t}_\rho$ would be normal if $\beta \neq 0$.

APPENDIX B

The Final Prediction Error Creterion

The FPE is defined as $\frac{1}{N} * \frac{(N+k)}{(N-k)} * SSR$, where k is the number of estimated coefficients. In other words the reduction of the estimated sum of squared residuals SSR has to be sufficiently large to outweigh the "penalty" of an increase in k by the addition of another lag. The search procedure for the minimum FPE in a system of equations can be described as follows: First, we will run univariate autoregressions for a variable y_1 and choose the lag length according to the minimum FPE (up to the eight-lag polynomial will be inspected). Second, add the second variable y_2 testing all lags. Causality runs from y_1 to y_2 if for any number of lags the FPE is smaller than in the univariate case for y_1. The same procedure will be followed in order to test for reversed causality i.e. if y_1 causes y_2. Finally, a third variable y_3 will be added to the optimal system of y_1 and y_2 to see whether y_3 causes y_1 or in the other case, causes y_2.

APPENDIX C

The Issue of Cointegration

Two series are defined as cointegrated when a linear combination of the two is stationary even though the series themselves are non-stationary. To illustrate the problem consider the simple model

$$y_{3t} = \delta y_{2t} + \varepsilon_t$$

where $\Delta\varepsilon_t = \rho\varepsilon_{t-1} + v_t$, v_t is white noise and y_3, y_2 are non-stationary series. The null hypothesis is taken to be non-cointegration $\rho = 0$. If δ were known, then a test for the null could be constructed along the lines of the Dickey-Fuller test, taking $y_{1t} = y_{3t} - \delta y_{2t}$ as the series that has a unit root under the null. However when δ is not known, it must be estimated from the data. But if the null $\rho = 0$ is true, δ is not identified. Thus, only if the series are cointegrated can δ be simply estimated by the cointegration regression. Engle and Granger (1987) propose several tests of the null of non-cointegration against the alternative of cointegration. Three of them are used in Table 4. The first test employs the DW statistic from the cointegrating regression. If the regression DW (CRDW) exceeds the respective critical value, then the null hypothesis is rejected, and the alternative cointegration one is accepted. The second and the third tests are the familiar Dickey-Fuller (DF) and augmented DF (ADF) ones, which are based on the following equations

$$\Delta\varepsilon_t = \rho\varepsilon_{t-1} + v_t \tag{C1}$$

$$\Delta\varepsilon_t = \rho\varepsilon_{t-1} + v_1\Delta\varepsilon_{t-1} + \dots + v_k\Delta\varepsilon_{t-k} + v_t \tag{C2}$$

These tests aim at investigating the existence of a unit root in the residuals of the cointegrating regression. The rejection of the unit root hypothesis implies that ε_t follows a stationary process, which means that y_2 and y_3 are cointegrated. In Table 4, under the headings DF and ADF are the t-values of the estimated ρ for the equations (C1) and (C2) respectively. The critical values of the above tests corresponding to the bivariate systems have been calculated by Engle and Granger (1987). For more than two cointegrating variables Engle and Yoo (1987) calculated critical values for the DF and ADF tests.

REFERENCES

AHMED E., ROSSER J.B., Jr. and SHEEHAN R.G., "A Comparison of National and International Aggregate Supply and Demand VAR Models: The USA, Japan and the EEG", *Weltwirtschaftliches Archiv.*, N. 2, 1989, 252-72.

BARRO R.J., *Macroeconomics*, 2nd edition, New York: John Wiley, 1987.

BLANCHARD O. and SUMMERS L., "Hysteresis in European Unemployment", *NBER Annual Supplement 1987*, 1986.

DICKEY D.A. and FULLER W.A., "Distribution of the Estimators for Autoregressive Time Series with a Unit Root", *Journal of American Statistical Association*, 1979, *74*, 427-31.

—— and ——, "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root", *Econometrica*, No. 4, 1981, *49*, 1057-72.

ENGLE R. and GRANGER C., "Co-integration and Error Correction: Representation Estimation and Testing", *Econometrica*, 1987, *55*, 251-76.

—— and YOO B., "Forecasting and Testing in Co-Integrated Systems", *Journal of Econometrics*, 1987, *35*, 143-59.

FISCHER S., "Recent Developments in Macroeconomics", *Economic Journal*, June 1988, *98*, 294-339.

FUCKLER J.S., "An Empirical Analysis of the Markets for Goods, Money and Credit", *Journal of Money Credit and Business*, N. 1, February 1985, *17*.

FULLER W.A., *Introduction to Statistical Time Series*, New York: John Wiley, 1976.

GOODHART C., "The Conduct of Monetary Policy", *Economic Journal*, June 1989, *99*, 293-346.

GRILLI V., "Seigniorage in Europe", in M. De Cecco and A. Giovannini, eds., *A European Central Bank?*, CEPR, 1988, 53-78.

HSIAO C., "Autoregressive Modelling and Money-Income Causality Detection", *Journal of Monetary Economics*, 1981, *7*, 85-106.

KING R.G. and PLOSSER C.I., "Money, Credit and Prices in a Real Business Cycle", *American Economic Review*, June 1984, *74*, 363-80.

KYDLAND F.E. and PRESCOTT E.C., "Time to Build and Aggregate Fluctuations", *Econometrica*, 1982, *50*, 1345-70.

LITTERMAN R.B. and WEISS L., "Money, Real Interest Rates, and Output: A Reinterpretation of Postwar US Data", *Econometrica*, 1985, *53*, 129-56.

LUCAS R.E., Jr., "Understanding Business Cycles", in K. Brunner and A. Meltzer, eds., *Stabilization of the Domestic and International Economy*, Carnegie-Rochester Conference Series, No. 1, 1977.

——, *Models of Business Cycles*, Oxford: Basil Blackwell, 1987.

MCCALLUM B.T., "On 'Real' and 'Sticky-Price' Theories of the Business Cycle", *Journal of Money, Credit, and Banking*, No. 4, 1986, *18*, 397-414.

NELSON C.R. and PLOSSER C.I., "Trends and Random Walks in Macroeconomic Time Series", *Journal of Monetary Economics*, 1982, *10*, 139-62.

PAGAN A.R. and WICKENS M.R., "A Survey of Some Recent Econometric Methods", *Economic Journal*, December 1989, *99*, 962-1025.

PERRON P., "Trends and Random Walks in Macroeconomic Time Series: Further Evidence from a New Approach", *Journal of Economic Dynamics and Control*, No. 2-3, 1988, *12*, 297-332.

SAID S.E. and DICKEY D.A., "Testing for Unit Roots in Auto-Regressive Moving Average Models of Unknown Order", *Biometrica*, 1984, *71*, 599-607.

SCHEIDE J., "On Real and Monetary Explanations of Business Cycles in West Germany", presented at the 4th Annual Conference of the EEA, Augsburg, W. Germany, 1989.

SIMS C.A., "Comparison of Interwar and Postwar Business Cycles: Monetarism Reconsidered", *American Economic Review*, 1980, *70*, 250-7.

——, "Policy Analysis with Econometric Models", *Brookings Papers in Economic Activity*, 1982, *1*, 107-52.

STULZ R.M. and WASSERFALLEN W., "Macroeconomic Time-Series Business Cycles, and Macroeconomic Policies", Carnegie-Rochester Conference Series on Public Policy, 1985, *22*, 9-54.

NELL'ECONOMIA DEL REGNO UNITO I CICLI ECONOMICI SONO REALI O MONETARI?

Questo articolo cerca di esaminare la validità empirica della teoria dei cicli economici reali per il Regno Unito. I risultati suggeriscono che la politica monetaria ha un ruolo primario nella spiegazione dei movimenti della produzione.

LITTERMAN R.B. and WEISS L., "Money, Real Interest Rates and Output: A Reinterpretation of Postwar US Data", *Econometrica*, 1985, 53, 129-56.

LUCAS R.E., Jr., "Understanding Business Cycles", in K. Brunner and A. Meltzer, eds., *Stabilization of the Domestic and International Economy*, Carnegie Rochester Conference Series, No. 5, 1977.

——, *Models of Business Cycles*, Oxford: Basil Blackwell, 1987.

MCCALLUM B.T., "On 'Real' and 'Sticky-Price' Theories of the Business Cycle", *Journal of Money, Credit, and Banking*, No. 4, 1986, 18, 397-414.

NELSON C.R. and PLOSSER C.I., "Trends and Random Walks in Macroeconomic Time Series", *Journal of Monetary Economics*, 1982, 10, 139-62.

PAGAN A.R. and WICKENS M.R., "A Survey of Some Recent Econometric Methods", *Economic Journal*, December 1989, 99, 962-1025.

PERRON P., "Trends and Random Walks in Macroeconomic Time Series: Further Evidence from a New Approach", *Journal of Economic Dynamics and Control*, No. 2/3, 1988, 12, 297-332.

SAID S.E. and DICKEY D.A., "Testing for Unit Roots in Auto-Regressive Moving-Average Models of Unknown Order", *Biometrica*, 1984, 71, 599-607.

SCHIPPS J., "On Real and Monetary Explanations of Business Cycles in West Germany", presented at the 4th Annual Conference of the EEA, Augsburg, W. Germany, 1989.

SIMS C.A., "Comparison of Interwar and Postwar Business Cycles: Monetarism Reconsidered", *American Economic Review*, 1980, 70, 250-7.

——, "Policy Analysis with Econometric Models", *Brookings Papers in Economic Activity*, 1982, 1, 107-52.

STULZ R.M. and WASSERFALLEN W., "Macroeconomic Time-Series Business Cycles and Macroeconomic Policies", Carnegie-Rochester Conference Series on Public Policy, 1985, 22, 9-54.

NELL'ECONOMIA DEL REGNO UNITO I CICLI ECONOMICI SONO REALI O MONETARI?

Questo articolo cerca di esaminare la validità empirica della teoria dei cicli economici reali per il Regno Unito. I risultati suggeriscono che la politica monetaria ha un ruolo primario nella spiegazione dei movimenti della produzione.

Rivista Internazionale di Scienze Economiche e Commerciali
Volume 39 (1992), N. 1, 69-89

DETERMINANTS OF HEALTH EXPENDITURES IN GREECE IN THE POSTWAR PERIOD: AN EMPIRICAL INVESTIGATION

by
George Karatzas *

I. *Introduction*

The purpose of this paper is to identify and assess the importance of the main factors affecting the level of health expenditures in Greece in the post-war period. The relative scarcity of the literature on the subject makes it useful to examine the determining factors of health expenditures for a small country in the postwar period in order to analyze these factors for policy making, since the lack of empirical research on the subject tends to hamper economic advice on national health expenditures.

With the publication of the Arrow (1963, 1967) and Grossman (1972b) studies, a new impetus has been provided for research on health expenditures. Kleiman (1974), and Newhouse (1977), among others, have examined the health expenditures of some western countries, whereas Pryor (1968) dealt with health expenditures in communist and noncommunist countries in general. Many of the available studies, however, being either normative or descriptive, are devoid of policy implications. Kleiman (1974) was one of the first to apply regression analysis in explaining national health expenditure changes. He was followed by the studies of Newhouse (1977), Cullis and West (1979), Council of Europe (1980), Maxwell (1981), O.E.C.D. (1985), Leu (1986), and Parkin, McGuire and Yule (1987) among others.

Before the 1970s economists focused their attention primarily on mi-

* Southern Connecticut State University, Economics Department, New Haven, Ct. (U.S.A.).
I wish to thank the editor of this journal and an anonymous referee for comments on an earlier draft of this paper.

croanalysis in dealing with health expenditures, whereby they more easily assessed costs, preferences and results, with special emphasis on cultural and institutional determinants of health expenditures (Glaser, 1970). Yet, policymakers at either the national or local level became equally concerned about the macro implications of health expenditures. Since the mid 1970s attention has been shifted to macro aspects of health expenditures. What was established before the 1970s in the various studies was the absence of a strong association between health-care expenditures and needs (ILO 1959; Abel-Smith 1967). Further, Abel-Smith argued that the relative share of health expenditures varies directly with the standard of living, as measured by the per capita income. Fraser (1972) on the other hand, argued that the ratio of health expenditures to total output is determined primarily by the ratio of government health expenditures to total health outlays, implying that nations that leave medical care to the private sector tend to demand more health services than those that have transferred this function to the public sector, and thus casting doubt on the fear that government-funded medicine opens up an uncontrollable flow of resources to health. In addition, Buchanan (1965) offered a theory that political processes would tend to curb activities channelled into the public sector more than if they were left in the private sector by arguing that if providers and patients are freed from financial constraints in their individual decisions, the sum of what benefits they all may wish to obtain as individuals will exceed what, as a nation, they wish to spend collectively. Thus, aggregate spending would be controlled through the political process more strictly than through the market.

But it was Newhouse who by 1977, after examining the relationship between health expenditures and income across countries reached the conclusion that factors other than income are most unlikely to explain variation in health expenditures, since approximately ninety percent of the variance of such expenditures can be explained by income, a view that was reasserted by an OECD study (1977). In this paper, however, we attempt to establish that in the case of Greece, in addition to income, other factors contributed to the determination of health expenditures.

Part I of this paper discusses the national health expenditures of Greece during the 1960-83 period. Part II consists of hypotheses that identify and explain the major determinants of the country's health expenditures. In Part III we test our hypotheses by applying the model to Greece and analyzing the results. Part IV summarizes our main conclusions.

Health expenditures constitute one of the largest areas of private and public expenditure growth. More specifically, nominal total health expenditures in Greece, that is both government and private, comprise around four percent of the country's nominal *GDP* (OECD, 1985) and over twelve percent of its nominal gross capital formation. Thus, the combined financial resources allocated to the health care sector of the country are small relative to the other western countries. Total health expenditures have been exhibiting a higher annual growth rate in real terms up to 1974 and a lower one thereafter, although in nominal terms they continued to grow because of the rapidly rising health care price index. After 1974 nominal government health expenditures have been rising at a higher rate than previously and their share in the country's nominal health expenditures have been increasing during the sample period.

Private health expenditures, in contrast, exhibited a constantly shrivelling rate of growth in real terms; this rate became negative after 1974. The decline in the rate of growth of private health expenditures in real terms has been mostly due to a presumptive bias since 1974 of a downward adjustment of the real level of private health expenditures, although in nominal terms their level has been continuously rising. The inflationary pressures which resulted from the 1973-74 economic turmoil and the international supply shocks reinforced the downward pressures on real private health expenditures, thus further contributing to the decline in the rate of growth of the country's real total health expenditures. Nominal ambulatory expenditures also declined after 1975, but resumed their ascendancy a year later, but at a low rate of growth. However, part of the observed rise in health expenditures could be attributed to what became to be known as the "accounting illusion", that is, the decline in the proportion of health care previously provided outside the market by family members or friends and hence used to go unrecorded, and which now is produced and sold in the market (Fuchs, 1972a; 1972b).

II. *Specification of the Empirical Model*

The model presented in this section includes three groups of independent variables, namely health stock variables, demographic variables, and economic variables as explanatory variables for the country's total health expenditures. Although cultural attitudes about health and family care, climate, and other environmental factors, and in general legal and institutional structures may affect or influence health expenditures, we contend that the variables included in the model far outweigh other influences. The first group of variables explain the supply factors, whereas the latter two emphasize demand for health expenditures [1]. However, in attempting to establish the hypotheses, we avoided the development of a strict theoretical model. Further, a simplified single-equation approach is used because of data availability and the high aggregation level of the data which requires that the model be estimated for annual aggregates.

We hypothesize that given the budget constraints, the total allocation of expenditures on health would be affected by a number of variables. The following group of variables are posited to influence national health expenditures. These variables are categorized as follows. First, the economic variables include: (1) the Gross Domestic Product at constant prices, (2) the health price index as expressed by the hospital care price index, (3) the income inequality as expressed by the share of wages in the Gross Domestic Product. Second, the health stock variables comprise: (1) the number of physicians, (2) the number of nurses, and (3) the number of hospital beds [2]. Third, we hypothesize that the demographic variables which tend to influence the level of national health expenditures include: (1) population changes, (2) population density, (3) population changes in cities with over one hundred thousand inhabitants, and (4) the total number of inhabitants under fifteen and over sixty-five years of age.

On the empirical level therefore, health expenditures can be expressed as a function of the economic, health stock, and demographic variables, with an appropriate time-lag adjustment for the health stock variables. The monocausal explanation that is advanced by Newhouse and his followers in their analysis of health expenditures cannot alone provide a plausible explanation of the determination process of health expenditures that prevailed in the country during the sample period. Thus, contrary to the view expressed by some investigators that non-income factors tend to have no influence on health expenditures, we contend that non-income variables are major determinants of health expenditures. We posit, therefore, that the main channel through which health expenditures are affected is through changes in the economic, health stock, and demographic variables, whereas any other factors played a minimal influence during this period.

The above hypotheses give rise to the following empirical specification of the total health expenditures regression equation in log-linear form.

[1] No attempt is made in this study to include political variables in the model since previous attempts by researchers to incorporate in their models the influence of political variables on health expenditures proved unsuccessful (Dye, 1966). Further, the establishment of a military dictatorship in the country in the late nineteen sixties and up to the mid nineteen seventies precludes the inclusion of political variables in the model.

[2] The number of midwives has been excluded from the health care stock variables, since it is small and does not appreciably influence the health stock variables.

$$In\,THE_t = a_1 + b_1\,In\,GDP_t - b_2\,In\,HPI_t + b_3\,In\,ID_t + b_4\,In\,POP_t + \\ + b_5\,In\,DENS_t + b_6\,In\,CIT_t + b_7\,In\,POU_t + b_8\,In\,DOC_{t-1} + \\ + b_9\,In\,NUR_{t-1} - b_{10}\,In\,BED_{t-1} + U$$

whereas the following variable definitions are employed:

THE = natural log of total health expenditures at constant prices.
HPI = natural log of health price index.
ID = natural log of income distribution as expressed by the ratio of nominal wages to nominal income.
DOC = natural log of total number of physicians.
NUR = natural log of total number of nurses.
BED = natural log of total number of hospital beds.
POP = natural log of total population.
DENS = natural log of population density.
CIT = natural log of cities with a population of over one hundred thousand inhabitants.
POU = natural log of total number of inhabitants under fifteen and over sixty-five years of age.
GDP = natural log of Gross Domestic Product at constant prices.
U = error term.

In summary our a priori expectations are:

$$b_1 > 0,\ b_2 < 0,\ b_3 > 0,\ b_4 > 0,\ b_5 > 0, \\ b_6 > 0,\ b_7 > 0,\ b_8 > 0,\ b_9 > 0,\ b_{10} < 0.$$

We contend that the income and price elasticities should be based on the level of real health expenditures rather than on the nominal health expenditures, since the health care consumer price index rose relatively faster in the postwar period than either the *GDP* deflator or the consumer price index. The nominal values of the health expenditures and of the *GDP* series are therefore stripped from inflationary pressures by expressing them at constant prices. By stripping both dependent and independent variables from inflationary pressures by different price indices, and thus expressing them in constant terms, the objections raised by Parkin (1987) that higher medical costs can account for the greater than one income elasticity are met.

We opted in favor of employing the hospital care price index as a deflator of the nominal health expenditures instead of the medical care and health services consumer price index, since the hospital care price index tends to be more representative of price trends in the health care sector.

The emergence of new products and procedures, coupled with the small sample size relative to the heterogeneity of the goods and services covered, precludes the health services consumer price index from accurately reflecting overall price trends in the health care sector. In addition, the health services consumer price index does not include hospital outlays or similar expenditures which are not directly related to the average consumer. Further, the health services consumer price series exclude certain expenditures, such for example as outlays on collective services and research and development, whereas as it has been documented, such downward presumptive biases are relatively smaller in the case of the hospital care price index for both private and government health expenditures (OECD, 1985, p. 43). Also, we experimented by constructing a medical price index by combining the pharmaceutical, the ambulatory, and the hospital price indices and weighing them according to their share in total health expenditures. The results of the experiment are not different from those reported in this paper. Further, it is expected on an a priori basis that changes in the *GDP* and greater income equality will be positively associated with the dependent variable, whereas a negative association between the price variable and the dependent variable is expected. In addition, one would expect a negative association between the hospital beds variable and the dependent variable, whereas for the remaining health stock variables and the dependent variable a positive association would be expected (Kleiman, 1974).

The *GDP* variable at constant prices is included in the equation by hypothesizing that changes in the country's *GDP* would tend to affect total health expenditures (Newhouse, 1977). In this study therefore we employ current income at constant prices, although current income suffers from two defects. First, it contains transitory income which tends to bias the estimated income elasticity, since the demand for medical services is associated with permanent income. Secondly, current income seems to be presumptively endogenous since the effect of sickness affects both the demand for medical services and income. But since there is no generally accepted method for measuring permanent income, current income at constant prices has been employed. Further, we hypothesize that the greater the inequality of income distribution, the lower is the level of total health expenditures. The ratio of total money wages to *GDP* is introduced as a proxy variable to capture the degree of income inequality [3]. Kleiman (1974) was the first to introduce a

[3] The share of labor in the *GDP* is not a perfect proxy variable for income inequality since income inequality is affected by the relative distribution of labor income itself. Also, the labor's share in the *GDP* can be a poor proxy variable for income inequality in a country

proxy variable to capture the degree of income inequality. The theoretical underpinning for such an undertaking was laid by Tobin (1970), who had argued that income inquality tends to be more intolerable for health-care services by society than for other goods or services. However, the income distribution variable, as well as other independent variables, have been excluded in the Newhouse (1977) model by making the dubious assumption that first, the per capita income variable takes care of all omitted independent variables, and second, that consumers possess homogeneous and of the same form utility functions (Parkin, 1987). The price variable is included in the equation by hypothesizing that there is a negative association between prices and real health expenditures. In the past prices have been excluded in some of the studies applying regression analysis on the ground that in most countries non-market rationing is prevalent (Newhouse, 1977). We consider however such a contention unjustifiable.

In addition to the economic variables, and given the budget constraints, it is assumed that contemporaneous or past augmentations in the country's health stock can affect real health expenditures. The total number of hospital beds, both public and private, with a year's time lag, is introduced in the equation as a proxy variable for the size of the country's health stock, although such a proxy variable does not capture all the effects of medicine [4]. In order to partly capture such effects, the total number of physicians and nurses, with a year's time lag, is also included in the equation [5]. It has been argued by Feldstein (1967), Detsky (1978), Evans (1974), Richardson (1981), and Wilensky and Rositer (1981) among others, that one of the most important determinants of health expenditures is the supply of physicians, since the physician has sufficient discretion in advising the patient and thus being able to offset changes in demand. We posit therefore that the above hypothesis is equally applicable in the case of Greece, and one would expect a positive association between health expenditures and the total number of actively engaged physicians. We hypothesize therefore that the number of physicians would be positively

---

where agricultural production is large, that is collinearity may be present between the labor's share and the degree of urbanization. But in the absence of a perfect proxy variable the share of labor to the country's *GDP* is employed.

[4] On an a priori basis one would expect that the stock of health care inherited from the past would be positively associated with income changes. This may result in collinearity between the income variable and the health-care stock variables.

[5] We also experimented by dividing the health-care stock variables by the country's population, but the obtained results were not significantly different from those presented on Table 1.

associated with the growth of real total health expenditures since the peculiar nature of demand for medical services assumes that the supplier creates his own demand by telling the patient what he must have. Thus, in this model we rationalize that the level of total health expenditures in any given period is influenced by the health stock of the previous period.

We hypothesize that causality runs from the health stock variable to health expenditures rather than the other way around as Roemer (1959, 1961) suggests. Our hypothesis is based on the argument that although it is possible that a direct positive relationship between bed availability and occupancy rate may exist, in a country like Greece an inverse relationship is more likely, since the country is facing a severe hospital bed shortage, that is, all available public hospital beds tend to be occupied, implying a negative relationship between the bed variable and health expenditures. In order to test this hypothesis we correlated occupancy rates and per capita beds. The correlation coefficient proved to be -0.84. The results of the correlation therefore indicate a negative relationship between occupancy rates and per capita beds, implying that the shorter stays are associated with higher costs per day, given the higher intensity of services which are performed at the beginning of the stay. In order to further test the hypothesis of higher intensity of service of health care, we correlated the number of nurses per hospital bed and hospital stay in days. The correlation coefficient is -0.80. The results of the correlation indicate the existence of a negative association between the number of nurses per bed and hospital stay in days, implying a greater intensity of service.

The demographic variables are introduced to capture the influence of population changes on real health expenditures and they encompass the effects of geographic distribution of population, the effects of population growth, the economies of scale resulting from the geographic distribution in the country's population, and the effects of changes in the age structure of the population. Data availability constrained both the time horizon and the number of explanatory variables chosen. It is assumed that population changes would be positively associated with the level of real health expenditures. Further, we hypothesize that population density would be positively related to the level of real health expenditures (Cochrane, 1978). The population density variable therefore is introduced to capture the adverse effect of urbanization on health expenditures.

The demographic variable indicating cities with a population over one hundred thousand inhabitants is included in the equation to account for the geographic distribution of the country's population (Leu, 1986). One would expect on an a priori basis that the higher the percentage of the population living in cities with over one hundred thousand inhabitants, the larger the economies of scale obtained in the provision of health care, and therefore the lower the level of real expenditures, as for instance in the case of occupancy of hospital beds. We encountered difficulties in applying the appropriate time series for this variable. Since statistics of cities with a population of over one hundred thousand inhabitants exist only on a five-year interval, it became necessary to interpolate the numbers in order to obtain annual series. It is assumed, in effect, that there has been a constant population growth in such cities during each five-year interval. The variable of the number of inhabitants under fifteen and over sixty-five years of age is included in the equation by hypothesizing that real health expenditures tend to be higher for the old and the young, compared to the rest of the country's population [6]. Thus, we posit that the changing demographic structures resulting from the aging of the population partly contribute to the growth of health expenditures (Emke-Poulopoulos, 1982; Madianos, 1988).

III. *Empirical Results*

Total Health Expenditures. – In order to empirically test the hypothesis we collected annual data from Greek and international sources, which are mentioned in the appendix of this study. The annual series are transformed into percentages to reduce serial correlation. Table 1 includes the estimated coefficients of the real total health expenditure functions, which yield the most satisfactory results on an a priori and statistical grounds. On the basis of the traditional statistical criteria the explanatory power of the regression planes is high. The results of the analysis are quite good in terms of the agreement between estimated and theoretically expected values of the coefficients, and possess the correct sign. The logarithmic form was chosen, even though in the arithmetic form the explanatory variables had the appropriate sign. Since the equations are double-log, the coefficients represent elasticities. The model empirically confirms the conventional wisdom of a strong linkage between the level of real total health expenditures and the three groups of the explanatory variables. Within the three groups of explanatory variables we encountered difficulties due to multicollinearity,

[6] In the case of the United States for instance (GIBSON, 1977) it has been estimated that those under fifteen and over sixty-five years of age account for over forty five percent of the total government expenditures.

TABLE 1

DETERMINANTS OF TOTAL HEALTH EXPENDITURES IN GREECE: 1961-1983

(1)

$$lnTHE_t = -34.469 - \underset{(-8.128)}{0.947\ lnHPI_t} - \underset{(-0.593)\,*}{0.267\ lnID_t} + \underset{(3.071)}{1.824\ lnDOC_{t-1}} + \underset{(1.346)\,*}{1.160\ lnCIT_t} + \underset{(2.40)}{6.703\ lnPOP_t}$$

$$R^2 = .968,\ \bar{R}^2 = .959,\ SEE = 0.042,\ DW = 1.33\ **$$

(2)

$$lnTHE_t = -17.719 + \underset{(4.415)}{0.874\ lnGDP_t} - \underset{(-5.767)}{0.513\ lnHPI_t} + \underset{(1.061)\,*}{0.421\ lnNUR_{t-1}} + \underset{(1.705)}{4.035\ lnPOP_t}$$

$$R^2 = .939,\ \bar{R}^2 = .925,\ SEE = 0.057,\ DW = 1.42\ **$$

(3)

$$lnTHE_t = -4.725 + \underset{(3.235)}{0.868\ lnGDP_t} - \underset{(-7.898)}{0.456\ lnHPI_t} + \underset{(1.464)\,*}{0.774\ lnNUR_{t-1}} + \underset{(0.983)\,*}{0.799\ lnDENS_t}$$

$$R^2 = .941,\ \bar{R}^2 = .929,\ SEE = 0.055,\ DW = 1.62\ **$$

(4)

$$lnTHE_t = -0.782 + \underset{(4.122)}{0.855\ lnGDP_t} - \underset{(-8.344)}{0.381\ lnHPI_t} + \underset{(1.727)}{0.669\ lnNUR_{t-1}}$$

$$R^2 = .929,\ \bar{R}^2 = .918,\ SEE = 0.059,\ DW = 1.44\ **$$

(5)

$$lnTHE_t = -3.652 + \underset{(3.121)}{0.759\ lnGDP_t} - \underset{(-5.036)}{0.620\ lnHPI_t} + \underset{(2.633)}{1.867\ lnDOC_{t-1}}$$

$$R^2 = .940,\ \bar{R}^2 = .930,\ SEE = 0.053,\ SW = 1.37\ **,\ \rho = 0.288$$

(6)

$$lnTHE_t = -29.476 + \underset{(2.430)}{0.446\ lnGDP_t} - \underset{(-6.962)}{0.919\ lnHPI_t} + \underset{(3.650)}{1.444\ lnDOC_{t-1}} + \underset{(3.320)}{5.760\ lnPOP_t}$$

$$R^2 = .962,\ \bar{R}^2 = .954,\ SEE = 0.044,\ DW = 1.19\ **$$

(7)

$$lnTHE_t = -25.162 - \underset{(-8.371)}{0.659\ lnHPI_t} - \underset{(-0.204)\,*}{0.156\ lnTHBED_{t-1}} + \underset{(7.040)}{3.068\ lnCIT_t} + \underset{(2.090)}{4.169\ lnPOP_t}$$

$$R^2 = .951,\ \bar{R}^2 = .941,\ SEE = 0.050,\ DW = 1.15\ **$$

continued TABLE 1

(8) $lnTHE_t = -61.174 - 0.867\ lnHPI_t - 0.070\ lnNUR_{t-1} + 5.113\ lnPOU_t$ $(-13.881)\quad(-0.221)\quad(6.919)$ $R^2 = .962,\ \bar{R}^2 = .956,\ SEE = 0.043,\ DW = 1.05$ **
(9) $lnTHE_t = 11.663 + 0.229\ lnGDP_t - 1.781\ lnID_t$ $(2.870)\quad(-3.841)$ $R^2 = .664,\ \bar{R}^2 = .631,\ SEE = 0.127,\ DW = 1.07$ **
(10) $lnTHE_t = -62.819 - 0.877\ lnHPI_t + 5.568\ lnPOU_t$ $(-19.941)\quad(21.564)$ $R^2 = .961,\ \bar{R}^2 = .958,\ SEE = 0.042,\ DW = 1.09$ **
(11) $lnTHE_t = -0.924 - 0.436\ lnHPI_t - 0.344\ lnID_t + 1.956\ lnNUR_{t-1}$ $(-5.879)\quad(-0.866)^*\quad(6.615)$ $R^2 = .871,\ \bar{R}^2 = .851,\ SEE = 0.080,\ DW = 1.26$ **
(12) $lnTHE_t = 12.109 - 0.548 - lnHPI_t$ (-3.498) $R^2 = .931,\ \bar{R}^2 = .924,\ SEE = 0.057,\ DW = 1.65,\ \rho = 0.976$
(13) $lnTHE_t = 4.701 + 0.057\ lnGDP_t$ $(0.090)^*$ $R^2 = .813,\ \bar{R}^2 = .795,\ SEE = 0.094,\ DW = 1.61,\ \rho = 0.851$
(14) $lnTHE_t = -4.804 + 1.116\ lnGDP_t$ (Nominal) (27.778) $R^2 = .998,\ \bar{R}^2 = .998,\ SEE = 0.039,\ DW = 1.79$

Note: The numbers in parentheses below the coefficients indicate *t*-values; R^2 = coefficient of determination; $\bar{R}^2$ = coefficient of determination adjusted for degrees of freedom; *D-W* = Durbin-Watson statistic; "ρ" = auto-regression coefficient estimated by using the Cochrane-Orcutt iterative technique to adjust for serial correlation; *S.E.E.* = standard error of estimate. * Insignificant at the 5% level. ** *D-W* in the indeterminate level.

that is, within each of the three groups of the explanatory variables collinearity proved to be present. Therefore, some of the explanatory variables are left out in order to overcome this problem.

In equation (1) two economic variables, one health stock variable, and two demographic variables have been included. All variables possess the expected sign, but the income distribution and one of the demographic variables are insignificant. The health stock variable, that is, the number of physicians, is significant and possesses a positive sign, confirming the hypothesis that physicians are in a position to influence the demand for medical services. The price variable is the most significant explanatory variable, and its negative coefficient is consistent with the negatively sloped demand curve, as it has been hypothesized. All the explanatory variables of the equation have coefficients greater than unity, with the exception of the economic variables, but the D-W statistic is in the indeterminate range, implying that no definite conclusion can be reached on the presence of serial correlation. Also, the results of the equation are quite good in terms of statistical fit. Further, the empirical results suggest that most of the variation of the dependent variable is absorbed by the contemporaneous values of the price, the lagged values of the health stock variable, and the contemporaneous demographic variable. Thus, the price, the health stock, and the demographic terms emerge as causally firm and account for a substantial fraction of the variance of real total health expenditures, suggesting that changes emanating from the values of price, the health stock, and demographic changes can exert substantial influence on health expenditures [7].

In equation (2) the income and price variables are included in the group of the economic variables, whereas from the demographic and health stock variables the total population and the total number of nurses are included among the independent variables. The *GDP* term is significant and smaller than one, suggesting that real health expenditures are income inelastic, and thus constitute a necessity rather than a luxury as Newhouse (1977, 1978) asserts. The smaller than one value of the income elasticity for real health expenditures can be explained by the fact that the nominal income has been deflated by the *GDP* deflator, whereas the nominal health expenditures have been deflated by the hospital care price index, which rose three times faster during the sample period than did the *GDP* deflator. Rising health costs can account for such rapid rise in the hospital care price index, a view which is also supported by the rapid rise of hospital expenses

[7] Also, we experimented by employing non-private health insurance as an explanatory variable, which proved to be insignificant. Non-private health insurance was significant only when ambulatory expenditures were employed as a dependent variable.

per patient day. The coefficient of the price variable also suggests that real total health expenditures are price inelastic. Further, the price term emerges as causally firm, and is the most significant variable in the equation accounting for a substantial fraction in the variance of health expenditures. The demographic variable is significant and positive, implying that part of the rise in the level of real health expenditures that occurred can be explained by changes that took place in the country's population. However, the health stock variable proved to be insignificant. What generally emerges from equation (2), is that the *GDP* variable is not the only determining factor in the variation of real total health expenditures and that real health expenditures are both income and price inelastic.

Equation (3) includes all the explanatory variables of equation (2), with the single exception of the population variable, which has been replaced by the population density. The economic variables proved to be significant, but the health stock and demographic variables are insignificant. The price variable, as in equation (1), proved to be the most significant explanatory variable. Equations (4) and (5) include two economic variables and one health stock variable. All coefficients are significant and possess the expected sign, with the price term being the most significant. The coefficients of the economic variables suggest that the real health expenditures are both price and income inelastic as it has been hypothesized, whereas the positive sign for the health stock variables suggests that the rise in the number of physicians and nurses during this period positively contributed to the growth of the real total health expenditures. Moreover, the physicians term in equation (5) is greater than one, implying that real total health expenditures are highly responsive to the changes in the number of physicians, confirming the hypothesis of physician-induced demand for health expenditures.

In equation (6) real total health expenditures were regressed on the two economic variables, that is on income and price, the health stock variable, and the demographic variable. All variables are significant and the coefficients possess the expected sign, with the population variable having the highest coefficient, suggesting a high responsiveness of real health expenditures to population changes. In equation (7), of the two economic variables which were included in equation (6), the *GDP* variable was dropped and only the price variable was retained. Also, a demographic variable was added to the regression equation, in addition to the health stock variable. The price term proved to be the most significant but its low coefficient suggests that real health expenditures are insensitive to price changes, whereas the population variable has a positive and significant coefficient. The total hospital beds variable has a negative coefficient, but is insignificant. The positive and significant coefficient of the cities with a population of over one hundred thousand inhabitants suggests that economies of scale did not apply to such cities during the sample period, a view corroborated by the findings of another study that employed a time series production model (Yfantopoulos, 1985, pp. 378, 400).

In equation (8) real total health expenditures are regressed on the price variable, the total number of nurses, and the population under fifteen and over sixty-five years of age. All the variables of the equation have the expected sign, with the exception of the health stock variable which is insignificant. The variable for the population under fifteen and over sixty-five years of age proved to be the most significant variable with the highest coefficient, and thus supporting the hypothesis that the young and the elderly account for a large fraction of the changes in the real health expenditures. In equation (9) two economic variables have been included, namely the real national income and the income distribution. Both variables proved to be significant. The income distribution variable is negatively correlated to the real total health expenditures, and has a coefficient greater than one, suggesting that real total health expenditures are sensitive to income distribution changes. Equation (10) includes the price, and the population under fifteen and over sixty-five years of age. Both of these variables are significant and have the expected signs, with the population under fifteen and over sixty-five years of age being the most significant and having a high coefficient, a view confirmed by the results of a survey (Madianos, 1988).

Equation (11) includes the income distribution, one economic variable, and one health stock variable. The price and health stock variables are significant and possess the expected sign, with the health stock variable having a coefficient greater than one. However, the income distribution variable is insignificant. In equation (12) real total health expenditures are regressed only on the price variable. The price variable has a coefficient of -0.548, suggesting that although increases in the prices in the health-care sector have a negative impact on real total health expenditures, such expenditures proved to be insensitive to price changes. In equation (13) real total health expenditures are regressed on the real *GDP*. The coefficient of the income variable is insignificant but smaller than one. However, in equation (14) where nominal total health expenditures are regressed on the nominal *GDP*, the coefficient of the *GDP* variable is positive and greater than one [8]. This seeming inconsistency between the real versus the nominal

[8] Also, we experimented by regressing real private health expenditures on real government

income elasticities could be explained by the rise in the share of nominal health expenditures in the country's nominal income, and the fall in the share of real health expenditures in the real *GDP* during the sample period. The rise in the share of the nominal health expenditures in the nominal *GDP* can be accounted for by the rapidly rising costs in the health sector, which in turn explains the greater than one value of the nominal income elasticity [9].

Generally, the income elasticity of health expenditures depends on the growth of the share of health expenditures with respect to the *GDP*. When the share of health expenditures to *GDP* is rising, the income elasticity is positive and greater than unity (Parkin, 1987, pp. 116). On theoretical grounds the above statement can be explained from the following: Let $THE = f(GDP)$, and THE/GDP be the share of the *GDP* spent on health.

$$\frac{d(THE/GDP)}{d(GDP)} = \frac{dTHE}{dGDP} \cdot \frac{1}{GDP} - \frac{THE}{GDP^2} =$$

$$= \frac{THE}{GDP^2} \cdot \left[\frac{dTHE}{dGDP} \cdot \frac{GDP}{THE} - 1 \right]$$

where $\frac{dTHE}{dGDP} \cdot \frac{GDP}{THE} = e_i$, the income elasticity of demand for health expenditures. Therefore, when $e_i < 1$ the share of health expenditures with respect to the *GDP* is declining, but when $e_i > 1$ the share is rising. However, it has been argued (Feldstein, 1981) that serious illness may lower family incomes, and thus cause a downward bias in the income elasticity. This may be the case in Greece since private medical insurance is almost nonexistent (Madianos, 1988), and private eleemosynary institutions do not participate in medical insurance.

In addition, these results suggest that income is not the only determi-

---

health expenditures. Real private health expenditures proved to be inversely related to real government health expenditures, and exhibit an elasticity of substitution smaller than unity, implying that households do not regard the health services provided by the government as good substitutes for those provided by the private sector. This view is corroborated through direct observation (Wood-Retsetakis, 1970; Madianos, 1988). The policy implication of such results is that improvements are needed in the provision of health services by the government.

[9] Also, we found through our experiments that different components of the real health expenditures have different income and price elasticities. Real hospital expenditures for instance, appear to be less sensitive to the real income changes relative to the real ambulatory expenditures.

ning factor of health expenditures, as claimed by Newhouse (1977, 1978, 1987). Also, the results of the estimated parameters of the health expenditures functions seem to provide support for the hypothesis that the economic variables, the health-care stock variables, and the demographic variables are the major determining factors of the country's national health expenditures during the sample period. This view is reinforced by the results obtained from the experiments conducted with various subsets of the explanatory variables. In addition, our elasticity estimates are most helpful for appraising the effect of variation in health expenditures. Further, our estimates of price and income elasticities fall within the range of such estimates reported for other countries, in particular those by Feldstein (1971, 1977), Fuchs and Kramer (1972), Davis and Russell (1972), and Rosett and Hung (1973) in the microeconomic, and Parkin (1987, 1988) in the macroeconomic sense. Also, the findings suggest that the economic, the health stock, and the demographic variables constitute key predictors for health expenditure behavior. One could also point out that the influence of the economic, the health stock, and the demographic variables transcends any political and institutional factors that might have exerted an influence on health expenditures during the sample period.

IV. *Conclusion*

We attempted in this paper to interpret changes in health expenditures in a small country during the postwar period by including three different causal phenomena, namely economic, health stock, and demographic variables in explaining changes in aggregate health expenditures. Our model seems to be the most appropriate model in explaining health expenditures, whereas other models are nested within it. In the determination of aggregate real health expenditures the real *GDP* variable proved to be significant and possesses the expected sign, and has a coefficient smaller than unity, implying that changes in the country's real *GDP* would tend to have a positive influence on the real aggregate health expenditures, but by a smaller percentage change than the real national income change. Although the *GDP* variable in most of the regression equations is highly significant, accounting for a high percentage of the variance in national health expenditures, this does not imply that the impact of non-income variables is minor. The price variable proved to be as important as the income variable in the determination of national health expenditures. When health expenditures are regressed on non-income variables some of such variables account for a high percentage of the variance in health expenditures. Of the four demographic variables the population term for over sixty-five and under fifteen years of age proved to be significant and its coefficient indicates that the quantitative effect is large, suggesting that the young and the elderly constitute the major consumers in health care.

Although we have attempted to establish that changes in the economic, health stock, and demographic variables play a leading role in influencing health expenditures in Greece during the period under study, other factors may have played a secondary role in the process. We contend however, that the three groups of variables which are included in our model outweigh other influences. Further, the high aggregation level of the data does not seem to invalidate the conclusion based on the results.

APPENDIX

Data, Definitions, and Sources

The data used in this paper are annual observations for the period 1961 to 1983 for Greece. The basic sources of data are (1) the National Statistical Service of Greece, *Statistical Yearbook of Greece*, (2) United Nations, *Demographic Yearbook*, (3) International Monetary Fund, *International Financial Statistics*, and (4) O.E.C.D. (1985). The discontinuity in some of the series necessitated a shorted sample period. The primary source for bed occupancy rates and hospital stay in days is the *Statistical Yearbook of Greece*, whereas the source for hospital expenses per patient day is O.E.C.D. (1985).

Variables

THE = total health expenditures at constant prices. This series is generated by dividing the nominal health expenditures by the hospital care price index. Source: O.E.C.D., (1985).

HPI = health care price index. Source: O.E.C.D. (1985).

DOC = total number of actively employed civilian physicians. Source: *Statistical Yearbook of Greece*, various issues.

NUR = total number of actively employed civilian nurses. Source: *Statistical Yearbook of Greece*, various issues.

BED = total number of government and private hospital beds. Source: *Statistical Yearbook of Greece*, various issues.

POP = the total population of the country. Source: *Statistical Yearbook of Greece*, various issues.

CIT = cities in Greece with over one hundred thousand inhabitants. Source: *Demographic Yearbook*, various issues.

GDP = Gross Domestic Product at constant prices. Source: International Monetary Fund, *International Financial Statistics Yearbook 1990*, line 99b.p, Washington, D.C., 1991.

DENS = population density. Source: *Demographic Yearbook*, various issues.

POU = total number of inhabitants over sixty-five and under fifteen years of age. Source: *Statistical Yearbook*, various issues.

ID = income distribution. This series is generated by dividing the nominal wages by the nominal Gross Domestic Product. Source: *Statistical Yearbook of Greece*, various issues.

REFERENCES

ABEL-SMITH B., *An International Study of Health Expenditures and its Relevance for Health Planning*, Geneva: World Health Organization, 1967.

ACTON J.P., "Nonmonetary Factors in the Demand for Medical Services: Some Empirical Evidence", *Journal of Political Economy*, 1975, *83*, 595-614.

ARROW K., "Uncertainty and the Welfare Economics of Medical Care", *American Economic Review*, December 1963, *53*, 941-73.

——, "Public and Private Values" in S. Hook, ed., *Human Values and Economic Policy*, New York: New York University Press, 1967.

BUCHANAN J., *The Inconsistencies of the National Health Service*, London: The Institute of Economic Affairs, 1965.

COCHRANE A.L., LEGER A.S. St., and MOORE F., "Health Service 'Input' and Mortality 'Output' in Developed Countries", *Journal of Epidemiology and Community Health*, 1978, *32*, 200-205.

COUNCIL OF EUROPE, *The Cost of Health Care in Member States of the Council of Europe and Finland*, Strasbourg: COE, 1980.

CROMWELL J. and MITCHELL J.B., "Physician-induced Demand for Surgery", *Journal of Health Economics*, No. 4, 1986, *5*, 293-314.

CULLIS J.G., "Medical Care Expenditures: A Cross-National Survey: An Extension to Developing Countries", University of Bath, mimeo, 1977.

—— and WEST P., *The Economics of Health*, Oxford: Martin Robertson, 1979.

DAVIS K. and RUSSELL L.B., "The Substitution of Hospital Outpatient Care for Inpatient Care", *Review of Economics and Statistics*, 1972, *54*, 109-120.

DETSKY A.S., *The Economic Foundations of National Health Policy*, Cambridge, Mass.: Ballinger, 1978.

DYE T.R., *Politics, Economics, and the Public: Policy Outcomes in the United States*, Chicago: Chicago University Press, 1966.

EMKE-POULOPOULOS I., "Le Vieillissement Démographique en Grèce", in H. Bartoli, ed., *Population, Travail, Chômage; Le Plein Emploi en Question: Etudes Coordonées par Henri Bartoli*, Paris: Economica, 1982, 183-230.

EVANS R.G., "Supplier-Induced Demand: Some Empirical Evidence and Implications", in M. Perlman, ed., *The Economics of Health and Medical Care*, New York: John Wiley, 1974, 162-73.

FELDSTEIN M.S., *Economic Analysis for Health Services Efficiency*, Amsterdam: North-Holland, 1967.

——, "Hospital Cost Inflation: A Study of Nonprofit Price Dynamics", *American Economic Review*, 1971, *61*, 853-72.

——, "Quality Change and the Demand for Hospital Care", *Econometrica*, 1977, *45*, 1681-1702.

——, *Hospital Costs and Health Insurance*, Cambridge, Mass.: Harvard University Press, 1981.

FRASER R.D., "Health and General Systems of Financing Health Care", *Medical Care*, July-August, 1972, *10*, 352-67.

——, "An International Study of Health and General Systems of Financing Health Care", *International Journal of Health Services*, N. 3, 1978, *3*, 369-96.

FUCHS V.R. (1972a), *Essays in the Economics of Health and Medical Care*, New York: National Bureau of Economic Research, Columbia University Press, 1972.

——(1972b), "The Growing Demand for Medical Care", in FUCHS (1972a, 61-68).

——(1972c), "The Basic Forces Influencing Costs of Medical Care", in FUCHS (1972a, 39-50).

——and KRAMER M.J., *Determinants of Expenditures for Physicians' Services in the United States, 1948-68*, New York: National Bureau of Economic Research, 1972.

GIBSON R.M., SMITH M. and FISHER C.R., "Age Difference in Health Care Spending: Fiscal Year 1976", *Social Security Bulletin*, August 1977, 3-14.

——and WALDO D.R., "National Health Expenditures in 1981", *Health Care Financing Review*, September 1982, 1-20.

GLASER W.A., *Paying the Doctor: Systems of Remuneration and their Effects*, Baltimore: Johns Hopkins University Press, 1970.

——, *Social Settings and Medical Organization: A Cross-national Study of the Hospital*, New York: Atherton Press, 1970.

GROSSMAN M. (1972a), "On the Concept of Health Capital and the Demand for Health", *Journal of Political Economy*, 1972, *80*, 223-55.

——(1972b), *The Demand for Health: A Theoretical and Empirical Investigation*, Occasional Paper No. 119, New York: National Bureau of Economic Research, 1972.

ILO (International Labor Office), *The Cost of Medical Care*, Geneva: ILO, 1959.

KLEIMAN E., "The Determinants of National Outlay on Health" in M. Perlman, ed., *The Economics of Health and Medical Care*, Proceedings of a Conference held by the International Economic Association in Tokyo, New York: John Wiley and Sons, 1974.

LEU R.E., "The Public-Private Mix and International Health Care Costs", in A.J. Culyer and Bengt Jonsson, eds., *Public and Private Health Services*, Oxford: Basil Blackwell, 1986.

MADIANOS M.G., ZARNARI O. et al., *Health and the Greek Society: An Empirical Investigation*, (in Greek) Athens: National Center for Social Research, 1988.

MAXWELL R.A., *Health and Wealth: An International Study of Health Care Spending*, Lexington, Mass.: Lexington Books, 1981.

NATIONAL STATISTICAL SERVICE OF GREECE, *Statistical Yearbook of Greece*, Athens: NSSG, 1986, various issues.

NEWHOUSE J.P., "New Estimates of Price and Income Elasticities of Medical Services", in R.N. Rosett, ed., *The Role of Health Insurance in the Health Services Sector*, Universities-NBER Conference Series No. 27, New York: Neal Watson, 1976, 261-312.

——, "Medical Care Expenditure: A Cross-National Survey", *Journal of Human Resources*, Winter 1977, 115-25.

——, *The Economics of Medical Care*, Reading, Mass.: Addison-Wesley, 1978.

——, "Cross-National Differences in Health Spending: What Do They Mean?", *Journal of Health Economics*, 1987, *6*, 159-62.

O.E.C.D., *Public Expenditure on Health*, OECD Studies in Resources Allocation, no. 4, Paris: O.E.C.D., 1977.

——, *Measuring Health Care 1960-1983: Expenditures Costs and Performance*, Paris: O.E.C.D., 1985.

—— *Financing and Delivering Health Care: A Comparative Analysis of the OECD Countries*, Paris: O.E.C.D., 1987.

PARKIN D., "Comparing Health Services Efficiency Across Countries", *Oxford Review of Economic Policy*, no. 1, 1988, *5*.

——, MCGUIRE A., and YULE B., "Aggregate Health Care Expenditures and National Income", *Journal of Health Economics*, 1987, *6*, 109-27.

PRYOR F.L., *Public Expenditures in Communist and Capitalist Nations*, Homewood, Ill.: Richard D. Irwin, 1968.

RICHARDSON J., "The Inducement Hypothesis: That Doctors Generate Demand for Their Own Services", in J. van der Gaag and M. Perlman, eds., *Health, Economics, and Health Economics*, Amsterdam: North-Holland, 1981, 189-24.

ROEMER M.I., "Bed Supply and Hospital Utilisation: A Natural Experiment", *Hospitals*, November 1961.

—— and SHAIN M., *Hospital Utilization under Insurance*, Chicago: American Hospital Association, 1959.

ROSETT R.N. and HUNG Lien-fu, "The Effects of Health Insurance on the Demand for Medical Care", *Journal of Political Economy*, March-April 1973, *81*, 281-305.

TOBIN J., "On Limiting the Domain of Inequality", *Journal of Law and Economics*, 1970, *13*, 263-78.

WILENSKY G.R., and ROSITER L.F., "The Magnitude and Determinants of Physician-Initiated Visits in the United States", in J. van der Gaag and M. Perlman, eds., *Health, Economics, and Health Economics*, Amsterdam: North-Holland, 1981, 215-44.

WOOD-RETSETAKIS A., *An Analysis of Health and Welfare Services in Greece*, Athens: Center for Planning and Economic Research, 1970.

UNITED NATIONS, *Demographic Yearbook*, New York: United Nations, 1985.

YFANTOPOULOS J., *Health Planning in Greece: Some Economic and Social Aspects*, (in Greek with a summary in English), Athens: National Center for Social Research, 1985.

LE DETERMINANTI DELLE SPESE SANITARIE IN GRECIA NEL PERIODO POSTBELLICO: UNO STUDIO EMPIRICO

Le spese sanitarie della Grecia nel periodo postbellico vengono confrontate con le possibili variabili esplicative. Si trova che le piú importanti variabili esplicative sono state le variabili economiche, quelle demografiche e gli stock di capitale sanitario. Nella determinazione delle spese sanitarie reali queste variabili si sono dimostrate significative, in particolare risultano esercitare una forte influenza sulla variabile dipendente le variabili non legate al reddito.

Wilensky G.R., and Rossiter L.F., "The Magnitude and Determinants of Physician-Initiated Visits in the United States", in J. van der Gaag and M. Perlman eds., *Health, Economics and Health Economics*, Amsterdam: North-Holland, 1981, 215-44.

Wood-Ritsatakis A., *An Analysis of Health and Welfare Services in Greece*, Athens: Center for Planning and Economic Research, 1970.

United Nations, *Demographic Yearbook*, New York: United Nations, 1982.

Yfantopoulos J., *Health Planning in Greece. Some Economic and Social Aspects* (in Greek with summary in English), Athens: National Center for Social Research, 1985.

LE DETERMINANTI DELLE SPESE SANITARIE IN GRECIA NEL PERIODO POSTBELLICO: UNO STUDIO EMPIRICO

Le spese sanitarie della Grecia nel periodo postbellico vengono confrontate con le possibili variabili esplicative. Si trova che le più importanti variabili esplicative sono state le variabili economiche, quelle demografiche e gli stocks di capitale sanitario. Nella determinazione delle spese sanitarie reali queste variabili si sono dimostrate significative, in particolare risultano esercitare una forte influenza sulla variabile dipendente le variabili non legate al reddito.

Rivista Internazionale di Scienze Economiche e Commerciali
Volume 39 (1992), N. 1, 91-94

FISCAL POLICIES AND GROWTH: AN EXTENSION

by
RICHARD J. CEBULA * and GERALD E. SCOTT *

1. *Introduction*

In a recent issue of *Public Choice*, Martin and Fardmanesh (1990) (hereafter, M & F) examine the impact of fiscal variables on economic growth. Using cross-sectional analysis, M & F derive policy implications regarding the impact of government spending, taxes, and deficits on the growth of real GNP. The purpose of this brief note is to extend the M & F analysis.

2. *Analysis*

M & F estimate a reduced-form equation using cross-section data for 76 countries. In the model, the growth in GNP is regressed against three policy variables:

(a) taxes, as a percent of *GDP*

(b) government expenditures, as a percent of *GDP*

(c) the government budget deficit, as a percent of *GDP*

The present paper extends and modifies the M & F analysis in a variety of ways, including the following:

– use of quarterly time-series data rather than cross-section data

– use of the tax rate rather than tax collections (which are endogenous)

– use of government purchases of goods and services rather than government expenditures (which are partly endogenous)

– allowing for the endogeneity of the budget deficit

* Department of Economics, Florida Atlantic University, Boca Raton, FL (U.S.A.).

– allowing for population size
– allowing for net exports
– allowing for monetary policy actions

The regression to be estimated here is given by:

$$CHPCY_t = a_0 + a_1 DEF_t + a_2 G_t + a_3 MAXT_t + a_4 M_t + a_5 NX_t + u \qquad (1)$$

where:

$CHPCY_t$ = the change in the per capita seasonally adjusted real *GNP* in quarter t;
a_0 = constant;
DEF_t = the seasonally adjusted federal budget deficit in quarter t, expressed as a percent of the seasonally adjusted trend *GNP* in quarter t;
G_t = the seasonally adjusted federal government purchases of goods and services in quarter t, expressed as a percent of the seasonally adjusted trend *GNP* in quarter t;
$MAXT_t$ = the maximum marginal federal personal income tax rate in quarter t, expressed as a percent;
M_t = the ratio of the average of the seasonally adjusted current quarter and preceding quarter net acquisitions of credit market instruments by the Federal Reserve System to the seasonally adjusted trend *GNP* in quarter t, expressed as a percent;
NX_t = the seasonally adjusted balance of trade in quarter t, expressed as a percent of the seasonally adjusted trend *GNP* in quarter t;
u = stochastic error term.

The model deals with quarterly data for the United States for the period 1957:1-1984:4. The seasonally adjusted middle-expansion trend *GNP* data were obtained from Holloway (1986, Table 2). The tax rate was obtained from the *Statistical Abstract of the United States*. The population, real *GNP*, government purchase, deficit and net export data were obtained from the *Economic Report of the President*. Finally, the open market operations data were obtained from the *Flow of Funds Acccounts* of the Federal Reserve System.

By expressing the growth variable ($CHPCY_t$) in *per capita* terms, our analysis allows for the impact of population size. In its specified form, variable G_t excludes all transfer payments and thus is treated as exogenous. Similarly, by defining the tax variable as a tax *rate*, we treat it as exogenous as well. Were the tax variable defined simply as tax *collections*, it would have to be treated as endogenous. By including M_t in the system, we allow for the impact of monetary policy; by including NX_t in the system, we allow for the fact that the United States is an open economic system. Finally, since the budget deficit is partly endogenous, its inclusion in the model introduces the possibility of simultaneous-equation bias. Accordingly, equation (1) is estimated using an instrumental variables technique (as well as the Cochrane-Orcutt procedure, to correct for first-order serial correlation), with the instrument being the one-quarter lag of the seasonally adjusted unemployment rate of the civilian labor force. The choice of instrument is based upon the fact that this unemployment rate systematically explains the budget deficit, whereas the lagged seasonally adjusted unemployment rate is not correlated with the contemporaneous error terms in the system.

The 2SLS estimate of equation (1) is given by:

$$CHPCY_t = 0.123 \underset{(-\ 2.33)}{-\ 0.36\ DEF_t} \underset{(+\ 1.72)}{+\ 0.24\ G_t}$$

$$\underset{(-\ 2.78)}{-\ 0.0008\ MAXT_t} \underset{(+\ 1.58)}{+\ 0.15\ M_t} \underset{(-\ 0.66)}{-\ 0.15\ NX_t},$$

$$DF = 105, \qquad DW = 1.60, \qquad Rho = 0.19 \tag{2}$$

where terms in parentheses are t-values.

In equation (2), the coefficient on G_t is positive but significant at only the eight percent level, providing only weak evidence that government purchases of goods and services act to elevate the value of $CHPCY_t$. On the other hand, the deficit is shown to exercise a *negative* and statistically significant impact on $CHPCY_t$. Similarly, the tax rate is also shown to exercise a *negative* and statistically significant impact upon $CHPCY_t$.

3. *Conclusion*

Equation (2) provides estimates of the impact of three fiscal variables upon the growth in per capita real *GNP* in the United States. To some degree, the choice of fiscal variables parallels that in M & F; nevertheless, as already noted above, there are a number of important differences between the M & F analysis and the present analysis. In particular, aside from defining economic growth in per capita terms (to allow for population size) and using time series data, the present note also defines G_t and the tax variable in exogenous terms, allows for endogeneity of the deficit, and allows for both monetary policy actions and openness of the United States economy. The primary conclusions regarding the impact of fiscal variables on per capita economic growth are:

(a) government purchases exercise only a weak impact;

(b) the budget deficit acts to significantly reduce the growth rate; and

(c) higher income tax rates significantly reduce the growth rate [1].

Clearly, over the long run, reduced government deficits (presumably to some degree accomplished by reducing outlays) and income tax rate cuts can be expected to yield major benefits in the United States.

REFERENCES

Holloway T., "The Cyclically Adjusted Federal Budget and Federal Debt: Revised and Updates Estimates", *Survey of Current Business*, 1986, *66*, 11-17.

Martin R. and Fardmanesh M., "Fiscal Variables and Growth: A Cross-sectional Analysis", *Public Choice*, 1990, *64*, 239-51.

POLITICHE FISCALI E CRESCITA

Questo articolo esamina empiricamente l'impatto della spesa statale, dell'imposizione fiscale e dei deficit di bilancio sulla crescita del PIL reale. L'articolo considera gli Stati Uniti e usa dati di serie temporali trimestrali per il periodo 1957-1984. I risultati principali sono:

– le spese statali hanno soltanto un impatto modesto

– il deficit statale ha un impatto significativo e negativo

– elevati saggi di imposte sul reddito riducono in modo significativo la crescita del PIL reale.

Il modello è stimato con variabili strumentali. I risultati implicano che può essere prudente ridurre il deficit di bilancio ma non attraverso un inasprimento fiscale. Un taglio alle spese statali sembra essere piú promettente.

[1] These results are consistent with those derived by M & F for 33 middle-income countries and 76 developed and developing countries; however, these results differ sharply from the M & F results for high-income countries per se.

LIBRI RICEVUTI (BOOKS RECEIVED)

NOTE ECONOMICHE: *Economia e finanza nell'ottica di un'Europa Unita*. 1990, Siena, Monte dei Paschi di Siena, pp. 528.

Nota editoriale. – Introduzione di P. Ciocca. – Sulla transizione dall'economia di comando al socialismo di mercato (*B. Jossa*). – Sequenza necessaria e credibilità della riforma nelle economie post-comuniste (*D.M. Nuti*). – La transizione delle economie dell'Est ed i mutamenti in atto negli equilibri economici europei (*V. Valli*). – La transizione finanziaria nei paesi dell'Europa Orientale (*M. De Cecco*). – Una sola Europa, ma con tre settori (*B. Gui*). – Cooperazione Est Ovest. La Comunità Europea tra competizione globale e bilateralismo (*P.C. Padoan*). – Trasferimenti verso i paesi dell'Europa Orientale e implicazioni per l'unificazione monetaria ed economica della Comunità Europea (*G. Basevi*). – Le relazioni finanziarie Est-Ovest: evoluzione e prospettive di sviluppo (*M. Mulino*). – L'Europa dell'Est e la Banca Europea per la Ricostruzione e lo Sviluppo. Un confronto con l'Europa del Piano Marshall (*C. Salleo*). – Il processo di trasformazione economica ed istituzionale nell'Est europeo: il caso polacco (*A. Chilosi*). – La perestroika tra disavanzo pubblico e stabilizzazione economica (*C. De Vincenti - A. Zevi*). – Perestroika e commercio con l'estero: quale futuro per il Comecon? (*C. Bentivogli*). – Partecipazione al commercio internazionale e convertibilità del rublo (*S. Trento*). – Le joint-venture in Unione Sovietica. Aspetti economici (*G. Graziani*). – La società mista nel diritto sovietico (*G. Crespi Reghizzi*).

RIVISTA DELLE SOCIETÀ: *La seconda direttiva CEE in materia societaria* (a cura di L. Buttaro e A. Patroni Griffi). 1984, Milano, Giuffré, pp. 800, L. 45.000.

Presentazione. – 1. Spese per la costituzione della società e generalità dei promotori (*G. Barbera*). – 2. La formazione del capitale sociale e l'art. 7 della seconda direttiva (*V. Manfredonia*). – 3. Osservazioni sulla liberazione delle azioni emesse a fronte di conferimenti in natura (*E. Lofoco*). – 4. La disciplina della valutazione dei conferimenti in natura nella previsione comunitaria e nell'art. 2343 c.c. (*G. Pellegrino*). – 5. Le acquisizioni di elementi patrimoniali appartenenti ai « soci » (*E. Lofoco, V. Manfredonia, G. Pellegrino*). – 6. Limiti alla distribuzione dell'attivo (*M. Castellano*). – 7. Limiti alla distribuzione di acconti sui dividendi nella disciplina comunitaria (*M. Lovecchio*). – 8. Sulla ripetibilità dell'utile irregolarmente distribuito (*M. Castellano*). – 9. Sottoscrizioni di azioni proprie (*L. Giorgio*). – 10. Acquisto di azioni proprie da parte della società emittente (*E. Sabatelli*). – 11. Acquisto di azioni proprie: finanziamento e contabilizzazione (*S. Fortunato*). – 12. Deroghe alla disciplina che regola l'acquisto di azioni proprie (*E. Sabatelli*). – 13. Acquisto di azioni proprie: dall'invalidità del contratto all'obbligo di alienazione (*A. Antonucci*). – 14. Sui diritti inerenti alle azioni proprie (*A. Antonucci*). – 15. Anticipazioni, prestiti o garanzie per l'acquisto di azioni proprie (*S.

Fortunato). – 16. Accettazione in garanzia di azioni proprie (*S. Fortunato*). – 17. Aumento di capitale e delega al consiglio di amministrazione (*F. La Malfa*). – 18. Obbligazioni convertibili in azioni (*F. La Malfa*). – 19. Aumento del capitale e obblighi dei sottoscrittori (*G. Tarantino*). – 20. Aumento di capitale non integralmente sottoscritto (*G. Tarantino*). – 21. Il diritto d'opzione nella disciplina comunitaria (*G. Labanca*). – Indice delle fonti normative. – Indice analitico-alfabetico. – Indice sommario.

RIVISTA DELLE SOCIETÀ: *L'informazione societaria*. Atti del Convegno Internazionale di Studi, Venezia, 5-6-7 novembre 1981. 1982, Milano, Giuffré, 2 voll., pp. 1424, L. 70.000.

1. INTRODUZIONE. I venticinque anni della Rivista (*A. Mignoli*). – 2. RELAZIONI GENERALI. L'informazione societaria e l'organo di controllo (*G. Rossi*). – L'informazione societaria e gli azionisti (*B. Visentini*). – 3. RELAZIONI NAZIONALI. L'information sociétaire en Belgique (*E. Wymeersch*). – L'information des associés en droit français (*R. Houin*). – Die Information der Gesellschafter und Dritter im deutschen Gesellschaftsrecht (*M. Lutter*). – Corporate Disclosure under Japanese Law (*M. Tatsuta*). – Corporate Disclosure in Great Britain (*C.M. Schmitthoff*). – L'information sociétaire aux Pays Bas (*C.W.A. Timmermans*). – La información social en España (*E. Verdera y Tuells*). – Corporate Disclosure in the United States (*L. Loss*). – Die Informationsrechte des Gesellschafters im Schweizerischen Recht (*P. Forstmoser*). – 4. TAVOLA ROTONDA. La società per azioni e le Commissioni di controllo. – 5. Comunicazioni. – 6. RELAZIONI DI SINTESI. Il punto di vista della Borsa (*G. Aloisio De Gaspari*). – Saluto (*M. Lutter*). – Sintesi generale (*G. Minervini*). – INDICE PER AUTORI.

ROSTOW W.W.: *Theorists of Economic Growth from David Hume to the Present. With a Perspective on the Next Century*. 1990, New York-Oxford, Oxford University Press, pp. 712, Lst. 32 (hardback).

Introduction. – PART I. Six Classical Economists: David Hume to Karl Marx. – PART II. Growth Theory Moves to the Periphery, 1870-1939. – PART III. Growth Analysis Post-1945: A Three-Ring Circus. – PART IV: Problems and Prospects. – Appendix: Models of Economic Growth.

STELLA Luciano: *Le teorie dell'aggiustamento della bilancia dei pagamenti*. 1990, Venezia, Marsilio Editori, pp. 196, L. 30.000.

Prefazione (di *A.P. Thirlwall*). – Introduzione. – La bilancia dei pagamenti. – L'aggiustamento della bilancia dei pagamenti tramite i prezzi in regime di cambi fissi: la teoria classica. – L'aggiustamento della bilancia dei pagamenti tramite la variazione del tasso di cambio: l'approccio delle elasticità. – L'aggiustamento della bilancia dei pagamenti tramite la variazione del reddito. – Effetti prezzo ed effetti reddito nell'aggiustamento della bilancia dei pagamenti. – L'approccio monetario alla bilancia dei pagamenti. – Il conflitto tra equilibrio esterno ed equilibrio interno. – Conclusioni. Bibliografia.